観測信号の分布に混合ガウス分布を、音声信号の振幅の事前分布に一般化ガンマ分布を用いたMMSE STSA 推定器による雑音低減

電気通信大学大学院 情報理工学研究科
総合情報学専攻 メディア情報学講座
吉田研究室所属

学籍番号 1130067
三田 太陽

主任指導教員　　西野哲朗　教授
指導教員　　吉田利信　教授

平成 25 年 1 月 30 日

要約

近年，スマートフォンやタブレット端末，カーナビゲーションシステムなどの普及により，音声認識の利用が広がっている．音声認識は，背景に雑音がほとんど無い場合には高い認識率が得られている．しかし実用的な環境では，雑音が混入してしまうために識率が著しく低下するという問題が残されている．雑音の混入に対処する方法を大別すると，認識システムを雑音に適応させる方法と，認識システムへの入力前に雑音を低減する方法がある．本研究では音声認識の前処理として音声信号中の雑音を低減することで，音声認識率の向上を目指す．

本研究では，音声の振幅の事前分布に一般化ガンマ分布を仮定し，観測信号の分布に混合ガウス分布を仮定して音声信号の推定値を計算する推定器 Parametric MMSE STSA GM を提案した．また観測信号の非音声区間からヒストグラムを作成し，混合ガウス分布との L^2 距離を最急降下法で最小化することで，観測信号の分布形状パラメータを推定する方法を提案した．

そして提案法である Parametric MMSE STSA GM と MMSE STSA [1]，Parametric MMSE STSA [2] との性能比較実験を行った．実験では，実用で混入が想定される雑音 5 種類と白色ガウス雑音を用い，評価尺度として単語正解率を用いた．その結果，空港雑音や店雑音では MMSE STSA および Parametric MMSE STSA と比べて単語正解率が向上した．レストラン雑音では，Parametric MMSE STSA よりも良く MMSE STSA と同等の単語正解率が得られた．白色雑音，バス雑音，オフィス雑音では，観測信号の分布形状パラメータの推定に失敗し，Parametric MMSE STSA よりも単語正解率が落ちたが，MMSE STSA と同等あるいはそれ以上の単語正解率であった．

今後の課題としては，分布形状パラメータの推定精度の向上が挙げられる．また音声信号の振幅の推定値は雑音の分散や事前 SNR によって決まるので，更なる性能向上を目指すには雑音推定や事前 SNR 推定などの性能を総合的に向上させる必要があると考えられる．

目 次

1 はじめに **1**

2 信号モデル **2**

2.1 観測信号 2

2.2 離散フーリエ変換と離散逆フーリエ変換 2

2.3 窓掛け 3

2.4 STFT 3

3 従来手法 **5**

3.1 MMSE STSA 5

3.2 Parametric MMSE STSA 8

4 Parametric MMSE STSA GM (提案法) **10**

4.1 原理 10

4.2 観測信号の分布形状パラメータの推定 11

4.2.1 3パラメータの場合 11

4.2.2 2パラメータの場合 12

5 評価実験 **14**

5.1 評価用信号 14

5.2 評価尺度 15

5.3 事前 SNR 推定 (Decision–Directed) 15

5.4 雑音推定 15

5.4.1 初期平均雑音推定法 15

5.4.2 重み付き雑音推定法 16

5.5 予備実験 1 17

5.5.1 実験条件 17

5.5.2 結果 17

5.5.3 考察 19

5.6 予備実験 2 20

5.6.1 実験条件 20

5.6.2 結果 20
5.6.3 考察 22
5.7 実験 24
5.7.1 実験条件 24
5.7.2 結果 24
5.7.3 考察 27

6 おわりに **27**

7 謝辞 **28**

付録 **30**

A 雑音信号のスペクトルコンポーネントの分布について **30**
A.1 実験 30
A.1.1 階級幅の決定方法 30
A.2 条件 30
A.3 結果 30
A.4 考察 33
A.5 まとめ 33

B 位相が認識率に与える影響 **34**
B.1 実験 1 34
B.1.1 条件 34
B.1.2 結果 34
B.2 実験 2 36
B.2.1 条件 36
B.2.2 結果 36
B.3 まとめ 37

C 振幅と位相の独立性 **38**
C.1 実験 38
C.2 実験条件 38
C.3 結果と考察 38

C.4 まとめ 38

D 音声の DFT コンポーネントにガウス分布を仮定することと振幅にレイリー分布を仮定することの等価性の証明 39

E Parametric MMSE STSA GM の導出 41

F 分布形状パラメータ r に関する制約条件の導出 46

1　はじめに

近年，スマートフォンやタブレット端末，カーナビゲーションシステムなどの普及により，音声認識の利用が広がっている．音声認識は，背景に雑音がほとんど無い場合には高い認識率が得られている．しかし実用的な環境では，雑音が混入してしまうため認識率が著しく低下するという問題が残されている．雑音の混入に対処する方法を大別すると，認識システムを雑音に適応させる方法と，認識システムへの入力前に雑音を低減する方法がある．本研究では音声認識の前処理として音声信号中の雑音を低減することで，音声認識率の向上を目指す．

雑音低減問題は，多くの場合「雑音が混入した音声信号から，音声の振幅スペクトルを推定する問題」として定式化されている．音声の振幅スペクトルを推定する最もシンプルな雑音低減法は，Spectral Subtruction 法 [3] (以降，SS 法) である．しかし，SS 法はミュージカルノイズと呼ばれる人工的な雑音を発生させてしまう．このミュージカルノイズは，音声認識率の改善を妨げるものとしてよく知られている．ミュージカルノイズが発生しにくく，Minimum Mean Square Error (以降，MMSE とする) に基づいた雑音低減法として，MMSE Short Time Spectral Amplitude [1] (以降，MMSE STSA とする) 推定器が提案されている．

MMSE STSA [1] では音声信号の振幅の事前分布にレイリー分布を，観測信号の分布にガウス分布を仮定し，音声信号の振幅スペクトルの推定値を計算する．Parametric MMSE STSA [2] では音声信号の振幅の事前分布がレイリー分布よりも左に歪んだ形状をすることから，音声信号の振幅の事前分布に一般化ガンマ分布を仮定し，音声信号の振幅スペクトルの推定値を計算する．本研究では，音声の振幅の事前分布は Parametric MMSE STSA と同様に一般化ガンマ分布を仮定するが，観測信号の分布には混合ガウス分布を仮定する．これらの仮定の下で，音声信号の振幅スペクトルの推定値を計算する推定器 Parametric MMSE STSA GM を提案する．また，観測信号の混合ガウス分布の形状パラメータを推定する方法を提案する．

2　信号モデル

本研究では，シングルマイクで観測された信号 (以降，観測信号) を周波数領域の信号に変換し，雑音低減を行う．

2.1　観測信号

ある離散時刻 n に観測された観測信号 y_n を式 (1) のように，音声信号 x_n と雑音信号 d_n の和として考える．

$$y_n = x_n + d_n \tag{1}$$

ただし音声と雑音はそれぞれ振幅の平均が 0 で互いに相関が無いと仮定する．

2.2　離散フーリエ変換と離散逆フーリエ変換

本研究では，時間領域の観測信号を周波数領域のスペクトル信号に変換する．長さ N の信号に対する離散フーリエ変換 (以降 DFT) と離散逆フーリエ変換 (以降 IDFT) は，それぞれ式 (2) と式 (3) で定義される．

$$X_k = \sum_{n=0}^{N-1} x_n \; w_n \; \mathrm{e}^{-j\frac{2\pi n}{N}k}, \quad (0 \leq k < N) \tag{2}$$

$$w_n \; x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k \; \mathrm{e}^{j\frac{2\pi n}{N}k}, \quad (0 \leq n < N) \tag{3}$$

ただし k は周波数ビン番号であり，w_n は窓関数である．窓関数 w_n について詳しくは 2.3 に述べる．

観測信号と音声信号，雑音信号に DFT を施し，それぞれ Y_k, X_k, D_k とすると，式 (1) は周波数領域で，式 (4) のように表現できる．この式 (4) を図示すると図 1 のようになる．

$$Y_k = X_k + D_k \tag{4}$$

また，これらを極座標で式 (5) のように表現する．

$$\begin{aligned} Y_k &= R_k \mathrm{e}^{j\vartheta_k} \\ X_k &= A_k \mathrm{e}^{j\alpha_k} \\ D_k &= N_k \mathrm{e}^{j\varphi_k} \end{aligned} \tag{5}$$

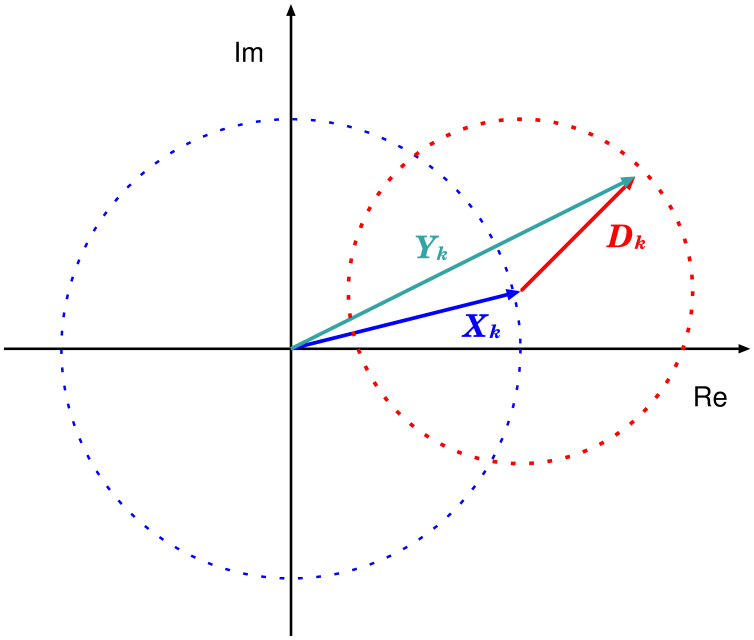

図 1: 観測信号：(青線) 音声信号 X，(赤線) 音声信号 D，(緑線) 音声信号 Y．点線は，音声信号の位相と雑音信号の位相の組合せによってそれぞれのベクトルが取り得る範囲を示す．

2.3 窓掛け

DFT では対象とする信号が周期的であることを前提としているので，観測信号にそのまま適用すると，本来存在しない周波数成分が表れてしまう．これを分析する信号に窓関数を掛けることで抑制する．音声信号を分析する際にはハミング窓やハニング窓を用いることが多い．図 2 にハミング窓とハニング窓，図 3 にその周波数特性を示した．図 3 から分るように，ハニング窓はハミング窓と比べてメインローブは太くなってしまうが，遠い周波数ビンでの減衰が −6dB/oct と大きい．そこで本研究では，遠い周波数ビンとの相関が少ないハニング窓を利用する．

2.4 STFT

本研究では，図 4 のように短時間毎に分析を行う短時間フーリエ変換 (Short Term Fourier Transform: STFT) を用いる．ただし，フレーム番号は省略する．なおフレームのシフト幅 (frame period) はフレーム幅 (frame length) の 50% として，再構成の際に足し合わせを行う Overlap–add 法を用いる．また実用を考えると計算量を抑える必要があるで，DFT の代りに高速フーリエ変換 (Fast Fourier Transform: FFT) を用いる．

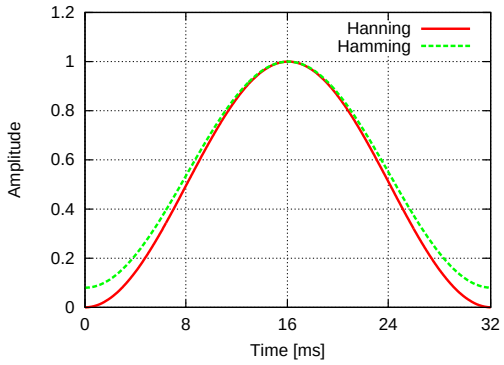

図 2: 窓関数：
(赤) ハニング窓，(緑) ハミング窓

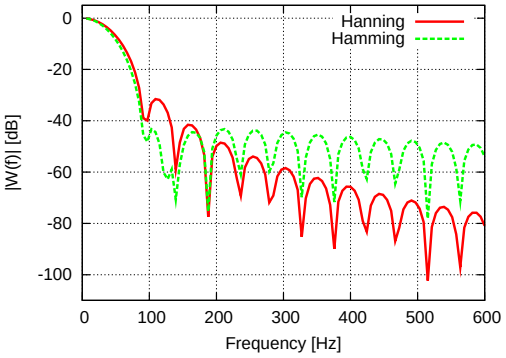

図 3: 窓関数の周波数特性：
(赤) ハニング窓，(緑) ハミング窓

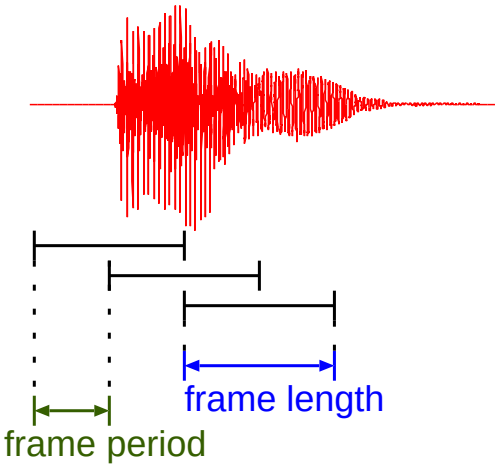

図 4: STFT 分析

3 従来手法

MMSE STSA (Minimum Mean Square Error Short Time Spectral Aamplitude) [1] は，雑音低減において有効とされている方法である．現在までに非常に多くの研究が行なわれ，スペクトルパワードメインの MMSE STSA 推定器 [5] や，MMSE LSA [6], $\beta - Order$ MMSE STSA [7]，Laplacian MMSE STSA [8]，など数多くの拡張が行われてきた．本節では，MMSE STSA とその拡張手法の 1 つである Parametric MMSE STSA [2] [9] について説明する．

3.1 MMSE STSA

MMSE STSA は，文献 [1] で Ephraim と Malah によって提案された音声信号の短時間スペクトル振幅を推定する推定器である．

MMSE STSA は，誤差関数

$$J = E\{(A_k - \hat{A}_k)^2 \mid Y_k\} \tag{6}$$

を最小化する音声信号の振幅の推定値 $\hat{A}_k$ を，式 (8) を用いて推定する．

$$\hat{A}_k = E\{A_k \mid Y_k\} \tag{7}$$

$$= \frac{\int_0^\infty \int_0^{2\pi} a_k p(a_k, \alpha_k) p(Y_k \mid a_k, \alpha_k)\ d\alpha_k\ da_k}{\int_0^\infty \int_0^{2\pi}\ p(a_k, \alpha_k) p(Y_k \mid a_k, \alpha_k)\ d\alpha_k\ da_k} \tag{8}$$

MMSE STSA では，音声および雑音のスペクトルコンポーネントの分布は，中心極限定理よりガウス分布であると仮定している．つまり音声信号の振幅の事前分布として式 (9) のレイリー分布を仮定し (付録 D 参照)，観測信号の分布として式 (11) のガウス分布を仮定している．そして音声の位相には式 (10) の一様分布を仮定している．

$$p(a_k) = \frac{2a_k}{\pi\lambda_x(k)} \exp\left\{-\frac{a_k^2}{\lambda_x(k)}\right\} \tag{9}$$

$$p(\alpha_k) = \frac{1}{2\pi} \tag{10}$$

$$p(Y_k \mid a_k, \alpha_k) = \frac{1}{\pi\lambda_d(k)} \exp\left\{-\frac{\mid Y_k - a_k \mathrm{e}^{j\alpha_k} \mid^2}{\lambda_d(k)}\right\} \tag{11}$$

ただし，

$$\lambda_x(k) \triangleq E\{\mid X_k \mid^2\}$$

$$\lambda_d(k) \triangleq E\{\mid D_k \mid^2\}$$

である．式 (9) のレイリー分布の分散を 1 にしたものを図 5 に，式 (10) の一様分布を図 6 に，式 (11) のガウス分布の平均を 0，分散を 1 にしたものを図 7 に示した．また音声の振幅と位相は独立であると仮定すると $p(a_k, \alpha_k) = p(a_k)p(\alpha_k)$ と表すことができる (付録 C 参照)．このとき式 (9),(10),(11) を式 (8) に代入して整理すると音声信号の振幅の推定値は，式 (12) で得られる．

$$\hat{A}_k^{\mathrm{MMSESTSA}} = \frac{\sqrt{\pi}}{2}\frac{\sqrt{v_k}}{\gamma_k}\ {}_1F_1(-0.5; 1; -v_k)\ R_k \tag{12}$$

ただし，

$$v_k \triangleq \frac{\xi_k}{1+\xi_k}\gamma_k$$
$$\xi_k \triangleq \frac{\lambda_x(k)}{\lambda_d(k)}$$
$$\gamma_k \triangleq \frac{R_k^2}{\lambda_d(k)}$$

であり，${}_1F_1(\alpha; \beta; \gamma)$ は合流型超幾何関数である．なお ξ_k は事前 SNR と呼ばれ，γ_k は事後 SNR と呼ばれる．

時間領域の音声信号の推定値は，推定した音声信号の振幅スペクトル $\hat{A}_k$ と観測信号の位相 ϑ_k を用いて式 (13) のように合成し (付録 B 参照)，$\hat{X}_k$ に IDFT を施すことで得る．

$$\hat{X}_k = \hat{A}_k \exp(j\vartheta_k) \tag{13}$$

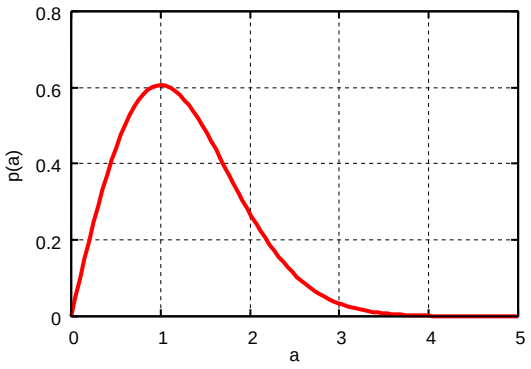

図 5: レイリー分布 (分散 1)

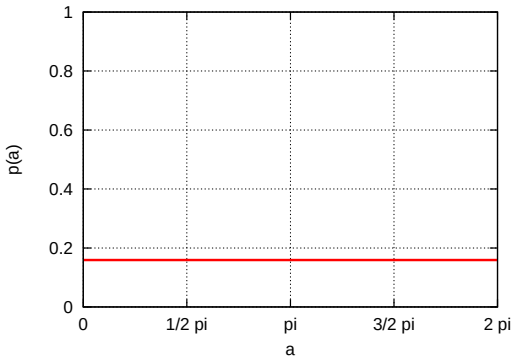

図 6: 一様分布

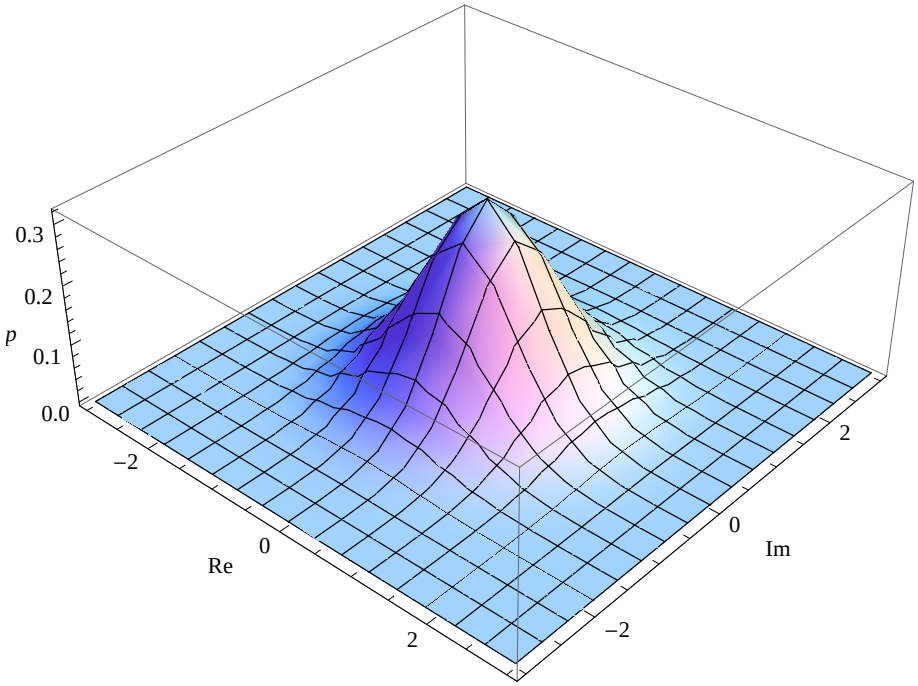

図 7: ガウス分布 (平均 0, 分散 1)

3.2 Parametric MMSE STSA

Parametric MMSE STSA は，文献 [2] で提案された振幅推定器である．3.1 節の MMSE STSA では，音声の振幅スペクトルの分布にレイリー分布を仮定していたが，実際の音声は図 8 のようにレイリー分布よりも左に歪んだ形状をしているため，音声の振幅スペクトルの分布を式 (14) の一般化ガンマ分布でモデル化する[1]．

$$p(a_k) = \frac{2\phi_k^{\eta}}{\Gamma(\eta)} a_k^{2\eta-1} \exp(-\phi_k a_k^2) \tag{14}$$

ただし η は分布の形状を決定するパラメータであり，$\phi_k = \eta/\lambda_x(k)$ かつ $0 < \eta \leq 1$ である．

式 (14),(10),(11) を式 (8) に代入して整理すると音声信号の振幅の推定値は，式 (15) で得られる．

$$\hat{A}_k^{\mathrm{PMMSESTSA}} = \frac{\sqrt{\tilde{v}_k}}{\gamma_k} \frac{\Gamma(\eta+0.5)}{\Gamma(\eta)} \frac{{}_1F_1(0.5-\eta;1;-\tilde{v}_k)}{{}_1F_1(1-\eta;1;-\tilde{v}_k)} \, R_k \tag{15}$$

ただし，

$$\tilde{v}_k \triangleq \frac{\xi_k}{\eta+\xi_k}\gamma_k \tag{16}$$

であり，$\Gamma(\cdot)$ はガンマ関数である．

[1] 文献 [2] では式 (14) をカイ分布と記載しており，文献 [9] では一般化ガンマ分布と記載している．本論文では，一般化ガンマ分布とする．

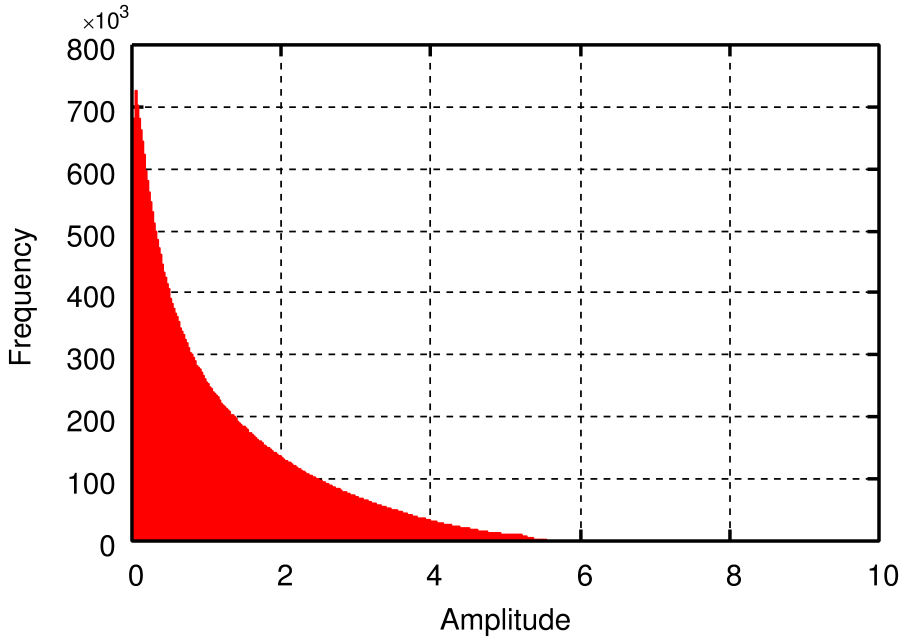

図 8: 音声の振幅のヒストグラム
(ATR セット B, 500 Hz–2 kHz, $\sigma^2 = 1$ となるように正規化してから作成)

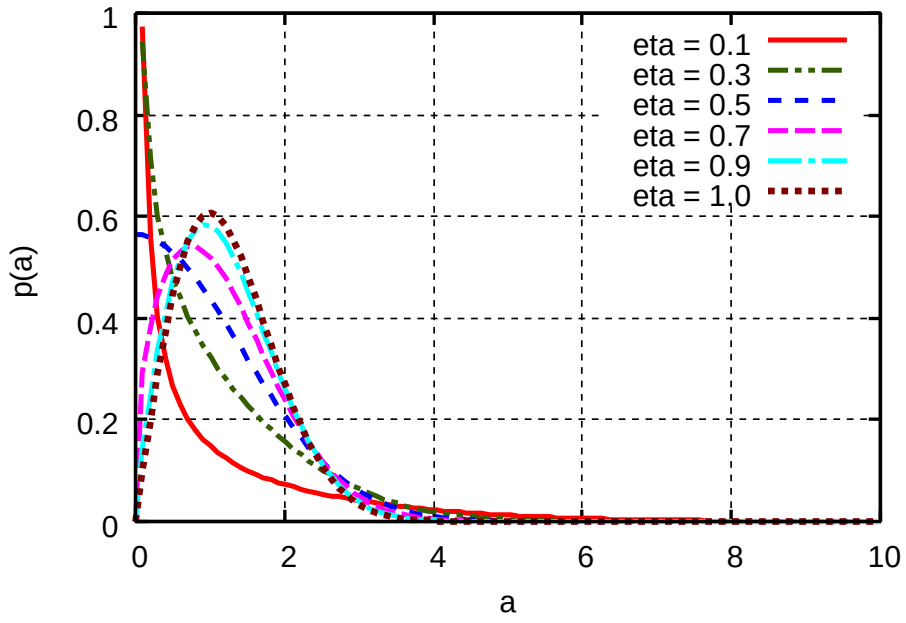

図 9: 一般化ガンマ分布 (分散 1)

4 Parametric MMSE STSA GM (提案法)

観測信号の分布に混合ガウス分布を，音声信号の事前分布に一般化ガンマ分布を仮定し，音声信号の推定値を計算する推定器を提案する．これを便宜上 Parametric MMSE STSA GM と呼ぶこととする．

4.1 原理

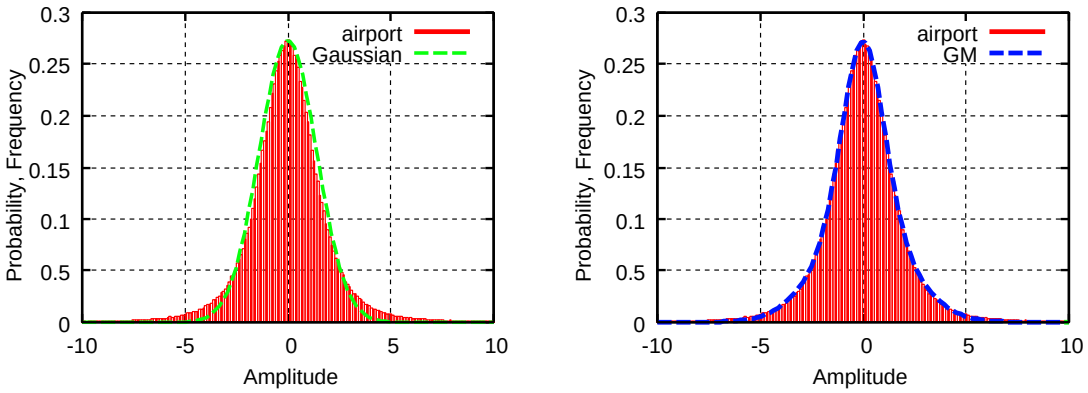

図 10: 空港雑音 D_k (500 Hz〜2 kHz, $\sigma^2 = 1$ となるように正規化) の実部の分布 (赤) とガウス分布 (左, 緑点線)，混合ガウス分布 (右, 青点線)

提案法では図 10 のように雑音の DFT スペクトルコンポーネント D_k の分布を分散 λ のみ k, i に依存し，$\boldsymbol{\theta}$ は k, i に依存しないものとして以下の混合ガウス分布で近似する (付録 A 参照).

$$Q(x, \lambda; \boldsymbol{\theta}) = \frac{r}{\pi z \lambda} \exp\left\{-\frac{|\, x\,|^2}{z\lambda}\right\} + \frac{1-r}{\pi w \lambda} \exp\left\{-\frac{|\, x\,|^2}{w\lambda}\right\} \tag{17}$$

ただし $r(0 \leq r \leq 1)$ は混合係数，$z(0 < z \leq 1)$, $w(w \geq 1)$ はそれぞれのガウス分布の広がりを調節する係数であり，$\boldsymbol{\theta} \triangleq [r, z, w]^T$ とする．

観測信号の分布はこれを用いると以下のように表される．

$$p(Y_k \mid a_k, \alpha_k) = Q(Y_k - a_k \mathrm{e}^{j\alpha_k}, \lambda_d(k); \boldsymbol{\theta}) \tag{18}$$

式 (18) と式 (14) を式 (8) に代入して整理すると，推定音声は式 (19) で得られる (付録 E

参照).

$$\hat{A}_k^{\mathrm{GM}} = \frac{r \cdot f_k(z, 0.5) + (1-r) \cdot f_k(w, 0.5)}{r \cdot f_k(z, 0) + (1-r) \cdot f_k(w, 0)} \tag{19}$$

$$f_k(\varsigma, \beta) = \frac{1}{\varsigma}\, \Gamma(\beta+\eta)\, \{\lambda_k(\eta, \varsigma)\}^{(\beta+\eta)} \cdot \exp\left(-\frac{R_k^2}{\varsigma \lambda_d(k)}\right) \cdot\ {}_1F_1\left(\beta+\eta; 1; v_k(\eta, \varsigma)\right) \tag{20}$$

ただし，

$$v_k(\eta, \varsigma) \triangleq \frac{\xi_k}{\eta\varsigma + \xi_k} \frac{\gamma_k}{\varsigma} \tag{21}$$

$$\frac{1}{\lambda_k(\eta, \varsigma)} \triangleq \frac{\eta}{\lambda_x(k)} + \frac{1}{\varsigma \lambda_d(k)} \tag{22}$$

である．

4.2 観測信号の分布形状パラメータの推定

雑音信号のスペクトル D_k の分散を 1 に正規化したものを x とし，そのヒストグラムを $P(x)$ とする．この $P(x)$ と式 (17) の混合ガウス分布 $Q(x, 1; \boldsymbol{\theta})$ との L^2 距離 [10]

$$g(\boldsymbol{\theta}) = \sum_x \left(P(x) - Q(x, 1; \boldsymbol{\theta})\right)^2 \tag{23}$$

を最小化することでパラメータ $\boldsymbol{\theta}$ の推定を行う．なお D_k のヒストグラム $P(x)$ は，観測信号の非音声部分から作成する．

この問題は，解析的に解くことができないので最急降下法を用いる．最小化する関数を $\boldsymbol{\theta}$ のそれぞれのパラメータで偏微分し，最急降下方向ベクトルを $\boldsymbol{d}$ とする．

式 (24) の更新式で分布形状パラメータの更新を行う．

$$\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n + \alpha \boldsymbol{d}^n \tag{24}$$

ただし n は更新回数であり，α は更新係数である．なお $\boldsymbol{\theta}^0$ は初期値として与え，$\boldsymbol{d}^n$ が十分小さくなったら更新を終了する．

4.2.1 3 パラメータの場合

求めたいパラメータを $\boldsymbol{\theta}_3 = [r, z, w]^T$ と表し，このときの最急降下方向ベクトルを式 (25) のように $\boldsymbol{d}_3$ と置く．

$$\boldsymbol{d}_3 = -\begin{bmatrix} \frac{\partial g(\boldsymbol{\theta}_3)}{\partial r} \\ \frac{\partial g(\boldsymbol{\theta}_3)}{\partial z} \\ \frac{\partial g(\boldsymbol{\theta}_3)}{\partial w} \end{bmatrix} \tag{25}$$

式 (23) と $\boldsymbol{\theta}_3$ を用いて $\frac{\partial g(\boldsymbol{\theta}_3)}{\partial r}$, $\frac{\partial g(\boldsymbol{\theta}_3)}{\partial z}$, $\frac{\partial g(\boldsymbol{\theta}_3)}{\partial w}$, 計算すると，それぞれ式 (26),(27),(28) のようになる.

$$\frac{\partial g(\boldsymbol{\theta}_3)}{\partial r} = \sum_x \left\{ -\frac{\mathrm{e}^{-\frac{x^2}{w\lambda}}}{\pi w\lambda} + \frac{\mathrm{e}^{-\frac{x^2}{z\lambda}}}{\pi z\lambda} \right\} (P(x) - Q(x, 1; \boldsymbol{\theta}_3)) \tag{26}$$

$$\frac{\partial g(\boldsymbol{\theta}_3)}{\partial z} = \sum_x \left\{ \frac{\mathrm{e}^{-\frac{x^2}{z\lambda}} r x^2}{\pi z^3\lambda^2} - \frac{\mathrm{e}^{-\frac{x^2}{z\lambda}} r}{\pi z^2\lambda} \right\} (P(x) - Q(x, 1; \boldsymbol{\theta}_3)) \tag{27}$$

$$\frac{\partial g(\boldsymbol{\theta}_3)}{\partial w} = \sum_x \left\{ \frac{\mathrm{e}^{-\frac{x^2}{w\lambda}} (1-r) x^2}{\pi w^3\lambda^2} - \frac{\mathrm{e}^{-\frac{x^2}{w\lambda}} (1-r)}{\pi w^2\lambda} \right\} (P(x) - Q(x, 1; \boldsymbol{\theta}_3)) \tag{28}$$

式 (26),(27),(28) を用いて，式 (24) の更新式で更新を行うことで，観測信号の分布形状パラメータを得る。

4.2.2 2 パラメータの場合

混合ガウス分布 $Q(x, \lambda_d; \boldsymbol{\theta})$ の 2 次のモーメントを計算すると，パラメータ r, z, w の間に式 (29) の関係式が成り立つ (付録 F 参照).

$$r = \frac{1-w}{z-w} \tag{29}$$

このとき r は式 (29) から求めるものとして，最急降下法で求めたいパラメータを式 (30) のように置く.

$$\boldsymbol{\theta}_2 = [z, w] \tag{30}$$

関数 $g(\boldsymbol{\theta}_2)$ を $\boldsymbol{\theta}_2$ のそれぞれのパラメータでの偏微分し，最急降下方向ベクトルを式 (31) のように $\boldsymbol{d}_2$ と置く.

$$\boldsymbol{d}_2 = - \begin{bmatrix} \frac{\partial g(\boldsymbol{\theta}_2)}{\partial z} \\ \frac{\partial g(\boldsymbol{\theta}_2)}{\partial w} \end{bmatrix} \tag{31}$$

式 (23) と $\boldsymbol{\theta}_2$ を用いて $\frac{\partial g(\boldsymbol{\theta}_2)}{\partial z}$, $\frac{\partial g(\boldsymbol{\theta}_2)}{\partial w}$, を計算すると，それぞれ式 (32),(33) のようになる.

$$\frac{\partial g(\boldsymbol{\theta}_2)}{\partial z} = \sum \left\{ \frac{\mathrm{e}^{-\frac{x^2}{z\lambda}}(-1+w)x^2}{\pi(w-z)z^3\lambda^2} - \frac{\mathrm{e}^{-\frac{x^2}{w\lambda}}(-1+w)}{\pi w(w-z)^2\lambda} \right.$$
$$\left. - \frac{\mathrm{e}^{-\frac{x^2}{z\lambda}}(-1+w)}{\pi(w-z)z^2\lambda} + \frac{\mathrm{e}^{-\frac{x^2}{z\lambda}}(-1+w)}{\pi(w-z)^2 z\lambda} \right\}$$
$$\cdot \left(P(x) - Q(x, 1; \boldsymbol{\theta}_2)\right) dx \tag{32}$$

$$\frac{\partial g(\boldsymbol{\theta}_2)}{\partial w} = \sum \left\{ \frac{\mathrm{e}^{-\frac{x^2}{w\lambda}}x^2\left(1 - \frac{-1+w}{w-z}\right)}{\pi w^3\lambda^2} + \frac{\mathrm{e}^{-\frac{x^2}{w\lambda}}\left(\frac{-1+w}{(w-z)^2} - \frac{1}{w-z}\right)}{\pi w\lambda} \right.$$
$$\left. - \frac{\mathrm{e}^{-\frac{x^2}{w\lambda}}\left(1 - \frac{-1+w}{w-z}\right)}{\pi w^2\lambda} - \frac{\mathrm{e}^{-\frac{x^2}{z\lambda}}(-1+w)}{\pi(w-z)^2 z\lambda} + \frac{\mathrm{e}^{-\frac{x^2}{z\lambda}}}{\pi(w-z)z\lambda} \right\}$$
$$\left(P(x) - Q(x, 1; \boldsymbol{\theta}_2)\right) dx \tag{33}$$

式 (32),(33) を用いて，式 (24) の更新式で更新を行うことで，観測信号の分布形状パラメータを得る。

5　評価実験

提案法の音声認識における有効性を調べるために，先行研究である MMSE STSA [1] と Parametric MMSE STSA [2] との性能比較実験を行う．また性能比較実験を行う前に，Parametric MMSE STSA における音声信号の分布形状パラメータ η を決定するための予備実験と提案法における観測信号の分布形状パラメータ推定法の予備実験を行う．

5.1　評価用信号

評価用信号として，音声は ATR 研究用日本語音声データベースセット A の音声を用い，雑音は NTT アドバンスドテクノロジ社の環境雑音データベース [11] より実環境雑音 5 種類と自作プログラムで生成した白色ガウス雑音を用いた。標本化周波数はすべて 16kHz にそろえて利用した。共通する実験条件を表 1 にまとめて示した．

表 1: 実験条件

音声	ATR 研究用日本語音声データベースセット A 計 281 単語 男女各 3 名 (faf, ffs, fym, mau, mht, mtk)
雑音	NTT アドバンスドテクノロジ社の環境雑音データベース [11] 実環境雑音 5 種類 (空港ロビー，オフィス，レストラン， バス車内，ショップ)，白色ガウス雑音
窓関数	ハニング窓
フレーム幅	512points (32ms)
フレームシフト幅	256points (16ms)
標本化周波数	16kHz
認識器	Viterbi デコーダ [12]
音響モデル	3 状態 4 混合の一方向性 HMM [12] (モノフォン)
特徴量	MFCC
高域強調係数	0.97

雑音重畳音声を，式 (34) のように雑音信号に重み係数 β を掛けて音声信号と足し合せることで作成した．なお重み係数 β は，式 (35) のように音声区間の音声信号パワーと雑音信号パワーから SNR が s [dB] になるように求めた．

$$y_n = x_n + \beta \times d_n \tag{34}$$

$$\beta = \frac{1}{10^{s/20}} \sqrt{\frac{\sum_{t=0}^{L} x_t^2}{\sum_{t=0}^{L} d_t^2}} \tag{35}$$

ただし L は音声区間のサンプル数であり，音声区間の取得にはコーパスに付属しているラベルの情報を用いた．

5.2 評価尺度

評価には，式 (36) で定義される単語正解率を用いた．

$$\text{単語正解率 [\%]} = \frac{\text{正解単語数}}{\text{全単語数}} \times 100 \tag{36}$$

5.3 事前 SNR 推定 (Decision–Directed)

STSA 推定器で音声の振幅スペクトルを推定するには，事前 SNR が必要である．しかし真の事前 SNR は得ることができないので，“Decision–Directed” 法 [1] を用いて観測信号から事前 SNR の推定値 $\hat{\xi}_k(i)$ を推定する．

$$\hat{\xi}_k(i) = \alpha \cdot \frac{\hat{A}^2_{k,i-1}}{\lambda_d(k, i-1)} + (1-\alpha) \cdot P[\gamma_k(i) - 1], \quad 0 \leqslant \alpha < 1 \tag{37}$$

ただし i はフレーム番号を表し，$P[\cdot]$ は式 (38) の半波整流を行う関数である．

$$P[x] = \begin{cases} x & \text{if x} \geqslant 0 \\ 0 & \text{otherwise} \end{cases} \tag{38}$$

5.4 雑音推定

事前 SNR と同様に，STSA 推定器で音声の振幅スペクトルを推定するには，雑音パワーが必要である．また前述の事前 SNR 推定にも雑音パワーが必要である．しかし真の雑音パワーは得ることができないので，雑音推定法を用いて観測信号から推定する．雑音推定方法として，初期平均雑音推定法，最小統計量に基く雑音推定 [13]，重み付き雑音推定 [14] などが提案されている．本研究では，初期平均雑音推定法と重み付き雑音推定法を用いる．

5.4.1 初期平均雑音推定法

初期平均法は，最もシンプルな雑音推定法である．$\hat{N}^2_k(i)$ を i 番目のフレーム，k 番目の周波数ビンでの雑音パワーの推定値とする．このとき初期平均推定における雑音パワーの推定値は式 (39) で得られる．

$$\hat{N}^2_k(i) = \begin{cases} \mid Y_{k,i} \mid^2 & \text{if } i < T_{IA} \\ \frac{1}{T_{IA}} \sum_{t=0}^{T_{IA}-1} \mid Y_{k,t} \mid^2 & \text{otherwise} \end{cases} \tag{39}$$

ただし T_{IA} は初期フレーム数であり，実験では T_{IA} は 7 frames とした．

5.4.2 重み付き雑音推定法

重み付き雑音推定法は，時間追従性の高い推定法である．以下に手順を示す．

1. あるフレーム i での事後 SNR を式 (40) で推定する．

$$\tilde{\gamma}_k(i) = \frac{|\ Y_{k,i}\ |^2}{\lambda_d(k, i-1)} \tag{40}$$

2. 式 (41) のように重み付き観測信号 $z_{k,i}$ を得る．重み関数 $W(\tilde{\gamma}_k(i))$ は式 (42) で定義される．図 11 に重み関数を示した．

$$z_{k,i} = W(\tilde{\gamma}_k(i))\cdot |\ Y_{k,i}\ |^2 \tag{41}$$

$$W(\tilde{\gamma}_k(i)) \quad = \begin{cases} 1 & \text{if } \tilde{\gamma}_k(i) < \tau_1 \\ 0 & \text{elseif } \tilde{\gamma}_k(i) > \tau_z \\ -\frac{1}{\tau_2-\tau_1}\cdot\tilde{\gamma}_k(i) + \frac{\tau_2}{\tau_2-\tau_1} & \text{otherwise} \end{cases} \tag{42}$$

ただし，τ_1, τ_2, τ_z は重み関数の形状を決定するパラメータである．

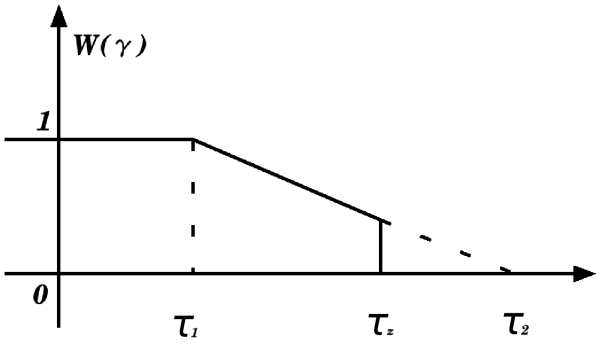

図 11: 重み関数

3. フレーム番号 i が T_{init} 未満であれば $|\ Y_{k,i}\ |^2$ を推定雑音パワーとする．
そうでなければ，現在のフレームから過去 L_z 個 (0 である要素を除く L_z 個を用いる) の重み付き観測信号 z の平均値を推定雑音パワーとする．

重み付き雑音推定のパラメータを表 2 に示した．なお初期フレーム数 T_{init} は初期平均法と同様に 7 frames とし，他のパラメータは文献 [14] で用いられたパラメータをそのまま用いた．

表 2: 重み付き雑音推定のパラメータ

τ_1	0 dB	τ_2	9 dB	τ_z	7 dB
T_{init}	7 frames	T_Z	20 frames		

5.5 予備実験 1

Parametric MMSE STSA [2] の最適な音声の分布形状パラメータを調べるために音声認識率による評価を行った.

5.5.1 実験条件

Parametric MMSE STSA [2] の音声の分布形状パラメータ η を $\eta = 0.1, 0.3, 0.5, 0.7, 0.9, 1.0, 1.5$ として単語正解率の調査を行った。なお分布形状パラメータ η の定義域は $0 < \eta \leq 1$ であるが，レストラン雑音では η の値が大きい方が認識率が良くなったので，$\eta = 1.5$ での評価も行った．また SNR は 0dB, 5dB, 10dB, 15dB とし，雑音推定法には初期平均法を用いた．その他の実験条件は表 1 と同様である.

5.5.2 結果

音声の分布形状パラメータ η を変えて雑音低減を行った音声の単語正解率を，図 12,13,14 に示した.

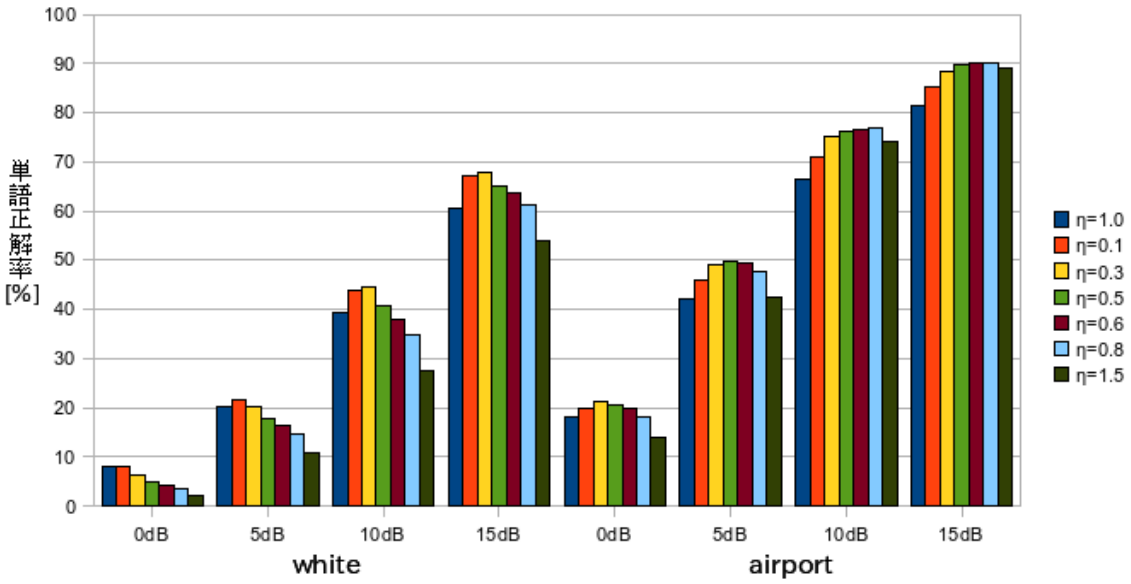

図 12: 単語正解率 (白色雑音，空港雑音)

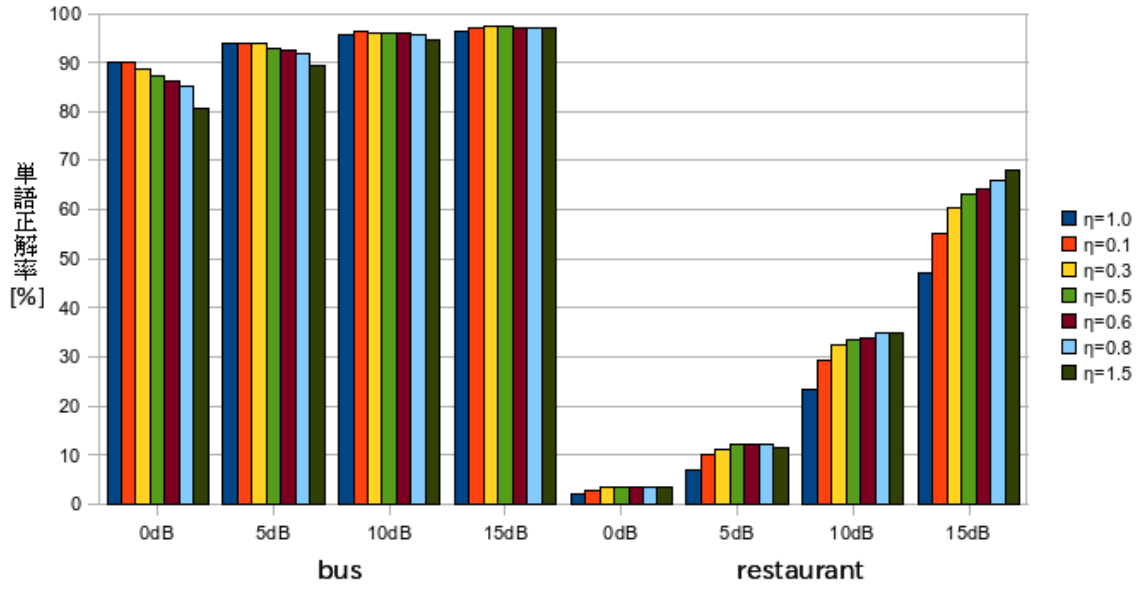

図 13: 単語正解率 (バス雑音，レストラン雑音)

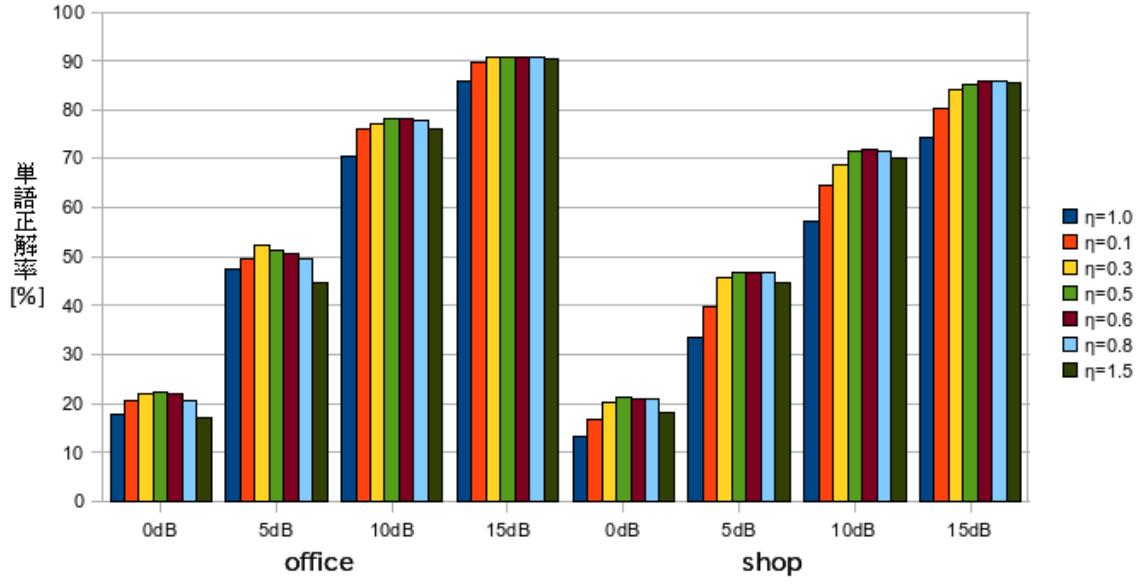

図 14: 単語正解率 (オフィス雑音，店雑音)

5.5.3 考察

全体的にみると，音声の分布形状パラメータ η が 0.3 と 0.5 のときに単語正解率が高くなった．しかしレストラン雑音では，η の値が大きい条件の方が単語正解率が良い結果になった．これは η の値を小さくして雑音低減量を増やしても，低減できずに残ってしまった雑音が本来の音声の特徴量を歪めてしまうためであると考えられる．

白色雑音の分布形状は，ガウス分布であることが明らかになっている．つまり実験に用いた 6 種類の雑音の中で，Parametric MMSE STSA の分布の仮定を最も良く満たす雑音である．この白色雑音では，$\eta = 0.3$ のときに単語正解率が最も良かった．単語正解率による評価では，音声の分布形状パラメータ η の最適値が雑音によって異なる結果となったが，雑音が異なっていても音声は同じなので音声の分布形状は変わらないはずである．そのため以降の実験では，雑音によらずに $\eta = 0.3$ を用いることにした．

5.6 予備実験 2

提案法である Parametric MMSE STSA GM の観測信号の分布形状パラメータの推定性能と最急降下法における分布形状パラメータの初期値の与え方についての調査を行った.

5.6.1 実験条件

以下の 4 通りの条件で実行した．なお SNR は-5dB, 0dB, 5dB, 10dB とし，雑音推定には重み付き雑音推定を用いた．また音声の分布形状パラメータ η は，5.5 節の予備実験の結果より 0.3 とした．その他の実験条件は，表 1 と同様である.

条件 1 [2params init–multi]
2 パラメータでパラメータ更新を行う。分布形状パラメータの初期値は $z = 0.1$〜1.0 の 0.1 毎, $w = 1.0$〜3.0 の 0.2 毎のすべての組み合わせで行い、L^2 距離が最も小さくなったパラメータを用いる。

条件 2 [3params init–multi]
3 パラメータでパラメータ更新を行う。
分布形状パラメータの初期値は $r = 0.1$〜1.0 の 0.1 毎, $z = 0.1$〜1.0 の 0.1 毎, $w = 1.0$〜3.0 の 0.2 毎のすべての組み合わせで行い、L^2 距離が最も小さくなったパラメータを用いる。

条件 3 [2params init–fixed]
2 パラメータでパラメータ更新を行う。
分布形状パラメータの初期値を $z = 1.0, w = 2.0$ とする。

条件 4 [3params init–fixed]
3 パラメータでパラメータ更新を行う。
分布形状パラメータの初期値を $r = 0.5, z = 1.0, w = 2.0$ とする。

5.6.2 結果

条件 1〜条件 4 の単語正解率を図 15〜図 17 に示した.

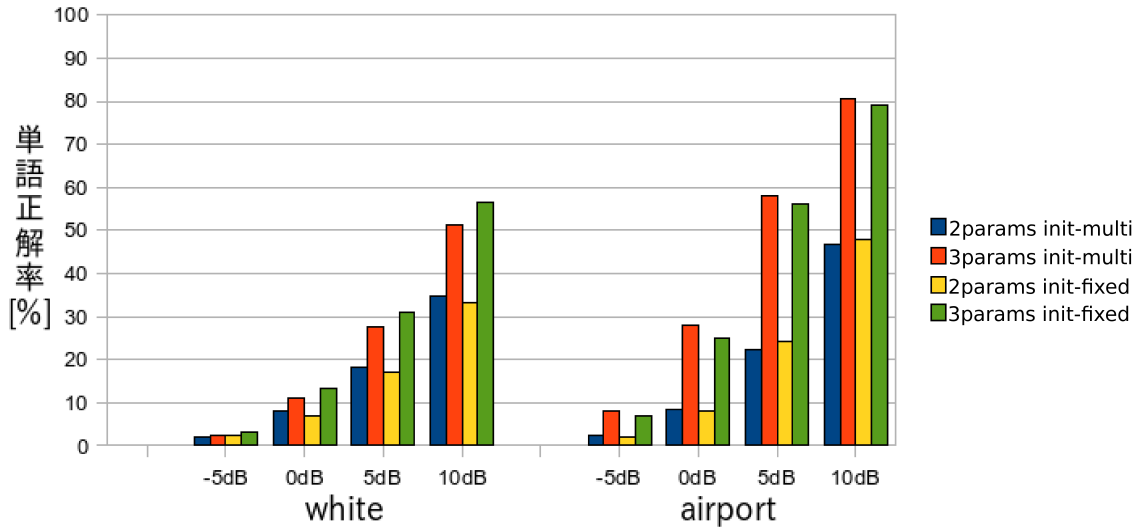

図 15: 単語正解率 (白色雑音，空港雑音)

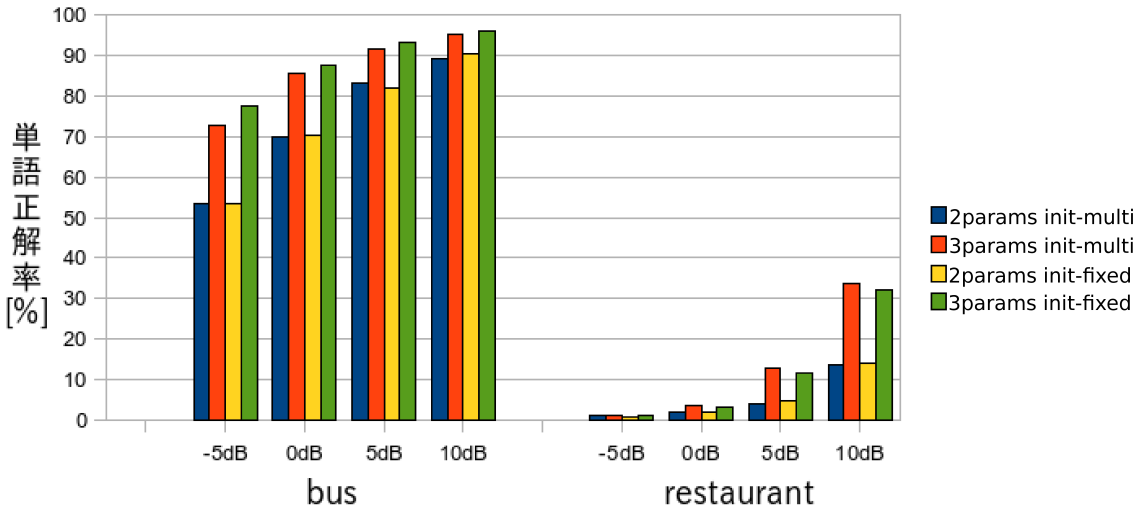

図 16: 単語正解率 (バス雑音，レストラン雑音)

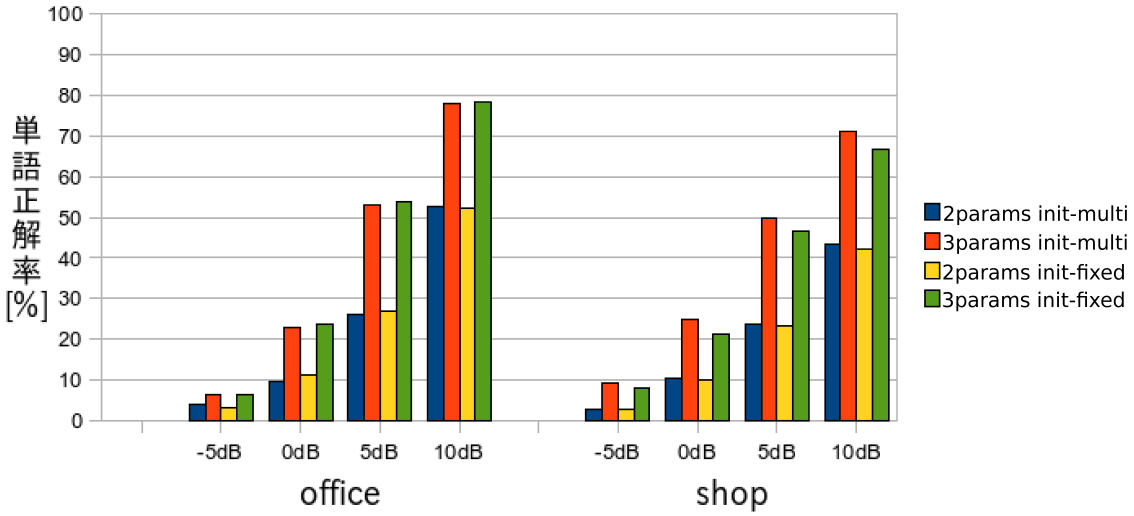

図 17: 単語正解率 (オフィス雑音，ショップ雑音)

5.6.3 考察

推定パラメータ数について 条件 1,3 [2params] の 2 パラメータの更新では，r の式に $z - w$ が分母に入っているので，$z - w$ の差が非常に小さい場合には r は発散してしまい，$0 < r \leq 1$ の条件を満さなくなってしまった．

条件 2,4 [3params] の 3 パラメータの更新では，r, z, w が全て定義域内に収まった。

初期値の与え方について 条件 1,2 [init–multi] では，ショップ雑音やレストラン雑音以外で条件 3,4 [init–fixed] よりも単語正解率が低い．これは音声ファイル先頭の短かい非音声区間から作成したヒストグラムと混合ガウス分布との L^2 距離が最も小さくなるパラメータを選択することから，雑音全体の特徴ではなく，短かい非音声区間だけの特徴を表現してしまっていると考えられる．

条件 3,4 [init–fixed] では，計算量が抑えられる利点が大きい．しかし空港雑音やレストラン雑音では，単語正解率が低くなってしまった．これは局所解に陥ってしまっている可能性が考えられる．

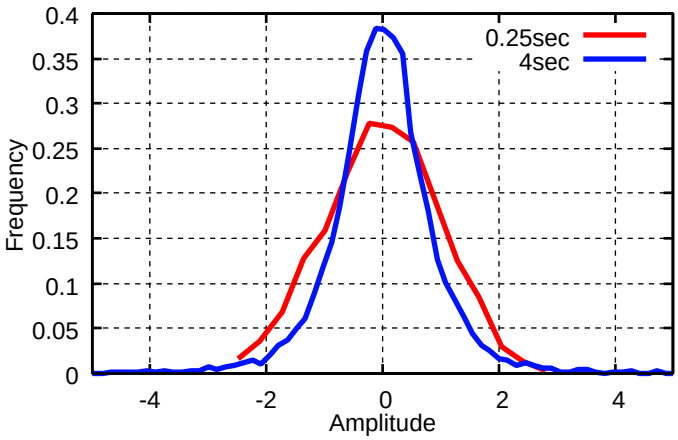

図 18: オフィス雑音のヒストグラム: (赤線) ファイル先頭から 0.25 秒間で作成，(青線) ファイル先頭から 4 秒間で作成

結論　初期値として実験的に決定したパラメータを幾つか与えたり，パラメータ選択の判定基準とする距離関数を変更したりすることで改善できるかもしれない．しかし Fig.18 を見ると，非音声区間で作成した雑音のヒストグラムから分布形状を推定するのは限界があると考えられる．

以上のことから 2 パラメータの更新では単語正解率が低く，アルゴリズムの安定性も低いことが分かった．そのため 3 パラメータの更新を利用することとする．初期値の与え方によって単語正解率が良くなる場合と悪くなる場合の両方があるので，5.7 節の実験では [init–fixed] と [init–mullti] の両方の条件を用いることにする．

5.7 実験

Parametric MMSE STSA GM と MMSE STSA [1]，Parmetric MMSE STSA [2] の雑音低減性能の比較実験を行った．

5.7.1 実験条件

以下の 6 つの条件で比較を行った。なお SNR は-5dB, 0dB, 5dB, 10dB とし，雑音推定には重み付き雑音推定を用いた．また条件 3〜5 における音声の分布形状パラメータ η は，5.5 節の予備実験の結果より 0.3 とした．さらに条件 4,5 における観測信号の分布形状パラメータの更新は 3 パラメータで行った．その他の実験条件は，表 1 と同様である．

条件 1 [NOISY] 雑音低減を行わない．

条件 2 [MMSE STSA] MMSE STSA [1] 推定器を用いて雑音低減を行う．

条件 3 [P-MMSE STSA] Parmetric MMSE STSA [2] 推定器を用いて雑音低減を行う．

条件 4 [GM init–multi] 提案法の推定器を用いて雑音低減を行う．
観測信号の分布形状パラメータの初期値は $r = 0.1$〜1.0 の 0.1 毎, $z = 0.1$〜1.0 の 0.1 毎, $w = 1.0$〜3.0 の 0.2 毎のすべての組み合わせで行い、L^2 距離が最も小さくなったパラメータを用いる。

条件 5 [GM init–fixed] 提案法の推定器を用いて雑音低減を行う．
観測信号の分布形状パラメータの初期値を $r = 0.5, z = 1.0, w = 2.0$ とする。

条件 6 [CLEAN] 雑音が混入していない音声．

5.7.2 結果

条件 1〜6 の単語正解率を図 19〜図 21 に示した．

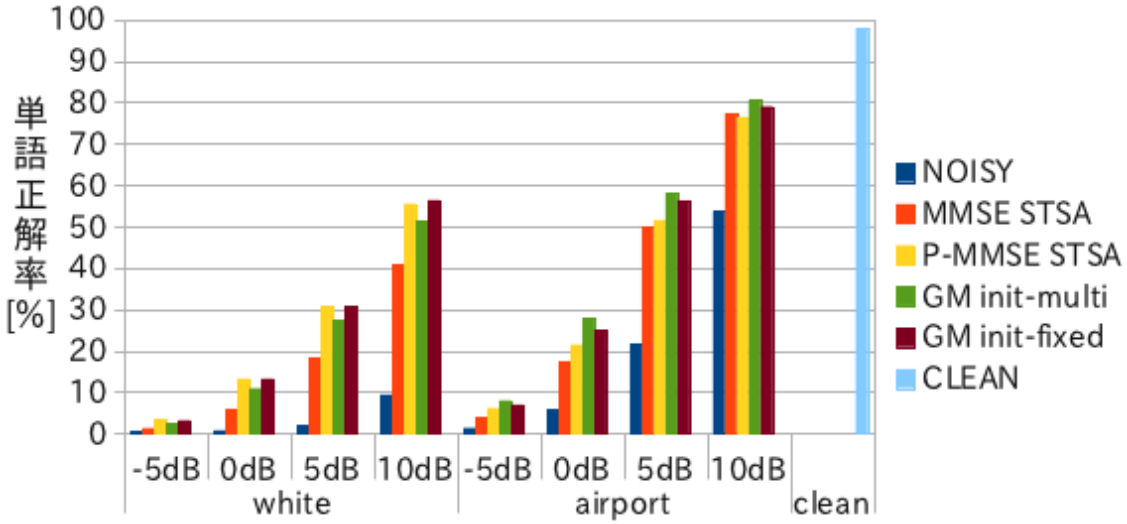

図 19: 単語正解率 (白色雑音，空港雑音)

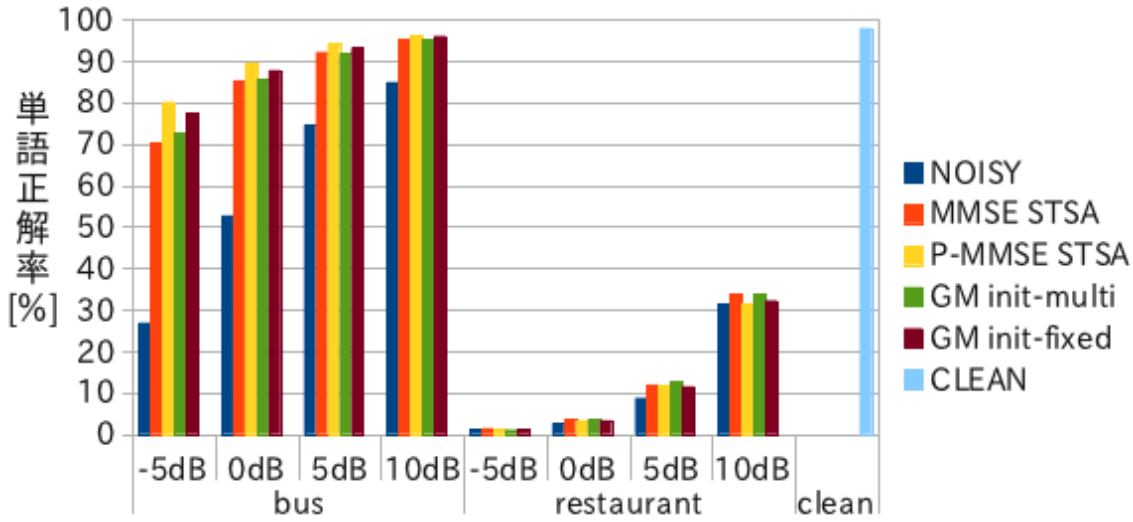

図 20: 単語正解率 (バス雑音，レストラン雑音)

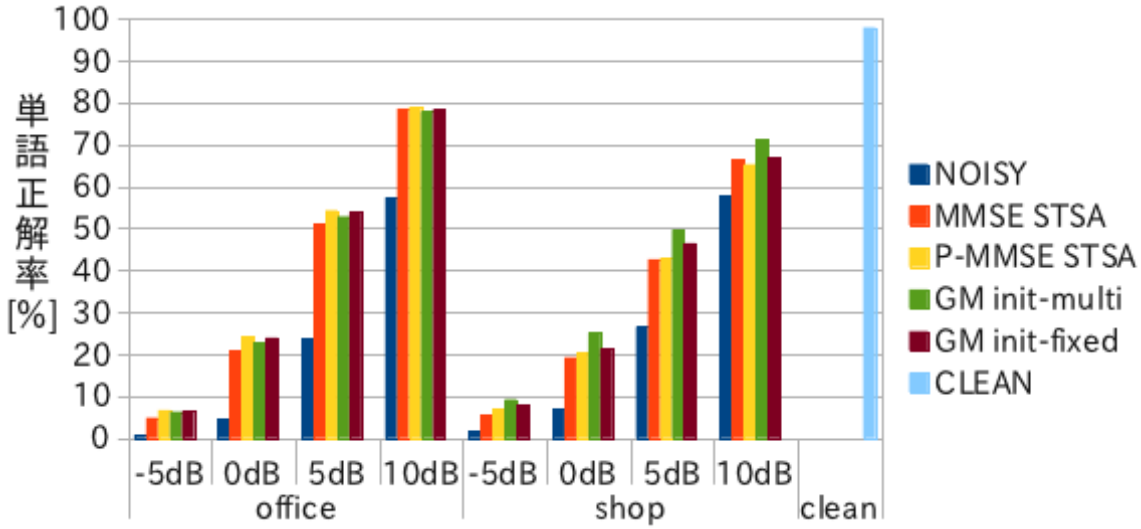

図 21: 単語正解率 (オフィス雑音，ショップ雑音)

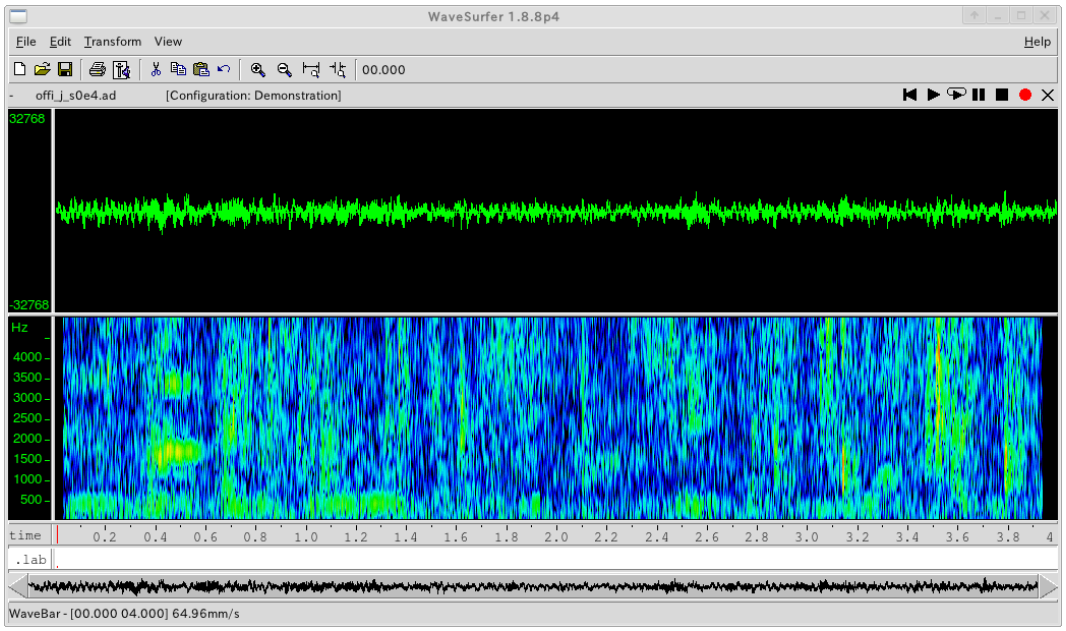

図 22: (上段) オフィス雑音の波形と (下段) そのスペクトログラム．
(Wavesurfer [15] を用いて表示)

5.7.3 考察

提案法は，空港雑音，店雑音，レストラン雑音において Parametric MMSE STSA よりも高い単語正解率が得られた．

しかし白色雑音では， Parametric MMSE STSA よりも単語正解率が下がってしまった．これは白色雑音のスペクトル D_k の分布形状がガウス分布であるにもかかわらず，分布形状パラメータの推定において本来の分布から遠いパラメータを推定してしまったためであると考えられる．バス雑音についても，雑音ファイルの先頭から 4 秒間のヒストグラムを見るとガウス分布に近い形をしていたため，白色雑音と同様に分布形状パラメータ推定に失敗したことが原因であると考えられる．

オフィス雑音では，図 22 に示すように非定常性が強く，音声が雑音として混入していることがわかる．このような雑音では，目的音声と雑音音声の区別ができないので，シングルマイクでの MMSE STSA による雑音低減の限界であると考えられる．

NOISY の単語正解率と雑音低減後の単語正解率を比べるとレストラン雑音を除いて大きな改善が得られた．レストラン雑音で大きな改善が得られなかったのは，雑音として混入している音声や突発的な雑音が本来の音声の特徴量を歪めてしまうためであると考えられる．また，クリーン音声の単語正解率と比較すると全体的に低い結果であった．

6 おわりに

本研究では音声認識の前処理として雑音低減を行い，認識率を向上させることを目的とし，Parametric MMSE STSA GM を提案した．また提案法である Parametric MMSE STSA GM における観測信号の分布形状パラメータの推定法を提案した．そして従来法である MMSE STSA [1] および Parametric MMSE STSA [2] との雑音低減性能の比較実験を行った．音声認識率による比較評価では，一部の雑音を除いて MMSE STSA および Parametric MMSE STSA よりも高い単語正解率が得られた．

今後の課題としては，分布形状パラメータの推定精度の向上が挙げられる．また音声信号の振幅の推定値は雑音の分散や事前 SNR によって決まるので，更なる性能向上を目指すには雑音推定や事前 SNR 推定などの性能を総合的に向上させる必要があると考えられる．

7　謝辞

研究を進めるにあたり，指導やアドバイスを頂きました西野哲朗教授，吉田利信教授，高木一幸助教授に心より感謝致します.

参考文献

[1] Y. Ephraim and D. Malah, “Speech enhancement using a minimum mean square error short-time spectral amplitude estimator,” IEEE Trans. on Acoust., Speech, Signal Processing, vol. ASSP-32, pp. 1109-1121, Dec. 1984.

[2] I.Andrianakis, P.R.White, “Speech spectral amplitude estimators using optimally shaped Gamma and Chi priors,” 2008 Elsevier.

[3] J.S.Lim, A.V.Oppenheim: Enhancement and band-width compression of noisy speech, Proc. IEEE 67(12), 1586–1604. 1979.

[4] Patric J. Wolfe, Simon J. Godsill, “Simple alternatives to the ephraim and malah suppression rule for speech enhancement,” IEEE 2001, 0–7803–7011-2/01.

[5] Patrick J.Wolfe and Simon J.Godsill, “Simple Alternatives to the Ephraim and Malah Suppression Rule for Speech Enhancement,” IEEE, 2001.

[6] Y.Ephraim, D.Malah, “ Speech Enhancement Using a Minimum Mean-Square Error Log–Spectral Amplitude Estimator,” IEEE Transactions on ACOUSTICS, speech,and signal processing, vol.ASSP–33,No2,APril 1985.

[7] ChangHuai You, SooNgee Koh, Susanto Rahardja, “Adaptive β–Order mmse stsa estimation for speech enhancement,” IEEE 2003, 0–7803–7663-3/03.

[8] Rainer Martin, “ Speech Enhancement Based on Minimum Mean-Square Error Estimation and Supergaussian Priors, ” IEEE Transactions on speech and audio processing, Vol.13, No.5, September 2005.

[9] 猿渡洋,” 音声信号処理における雑音抑圧技術の最新動向,” 信学技報, IEICE Technical Report, SP2010–103 (2011-01).

[10] 東京工業大学　杉山　将, “確率分布間の距離推定,” http://sugiyama-www.cs.titech.ac.jp/ sugi/2007/Canon-MachineLearning28-jp.pdf, アクセス (2012/01/3).

[11] NTT アドバンスドテクノロジ株式会社, “Ambient Noise Database for Telephonometry 1996” 取扱説明書 (1996).

[12] Yutaka TSUBOI,Takehiro IHARA,Kazuyuki TAKAGI,Kazuhiko OZEKI, “The Use of Overlapped Sub-Bands in Multi-Band, Multi-SNR, Multi-Path Recognition of Noisy Word Utterances,” IEICE TRANSACTIONS on Information and Systems Vol.E91-D No.6, pp.1774–1782.

[13] Rainer Martin, “ Spectral Subtraction Based on Minimum Statistics, ”Proc. EUSIPCO ’94, pp. 1182–85, Sept. 1994.

[14] 加藤正徳, 杉山昭彦, 芹沢昌広,“ 重み付き雑音推定と MMSE STSA 法に基づく高音質雑音抑圧, ” 電子情報通信学会論文誌 A, Vol.J87–A No.7, pp.851–860, July. 2004.

[15] Centre for Speech Technology at KTH, “ TMH KTH : WaveSurfer, ” [http://www.speech.kth.se/wavesurfer/], アクセス (2013/1/28).

[16] Scott, David W. “On optimal and data-based histograms.” Biometrika 66 (3): 605–610. (1979).

[17] I.S. Gradshteyn and I.M. Ryzhik ; incorporating the 4th ed. prepared by Yu. V. Geronimus, M. Yu. Tseytlin ; translated from the Russian by Scripta Technica, inc., and edited by Alan Jeffrey Corr. and enl. ed. / prepared by Alan Jeffrey San Diego ; Tokyo : Academic Press, 1980.

付録

A　雑音信号のスペクトルコンポーネントの分布について

観測信号の分布について，実際の信号における統計的な特徴を確認するために，雑音信号のスペクトルコンポーネントの分布ついて調査する．

A.1　実験

代表的な特定環境の雑音 5 種類と人工的に生成した白色雑音についてスペクトルコンポーネントの分布を調査した．

A.1.1　階級幅の決定方法

階級幅 h は，式 (43) の Scott's Choice [16] を用いて決定した．

$$h = \frac{3.5\sigma}{n^{1/3}} \tag{43}$$

ただし σ はサンプルの標準偏差であり，n はサンプル数である．

A.2　条件

ヒストグラムは，周波数ビン毎の標準偏差で割って正規化した．その他の実験条件を表 3 に示した．

表 3: 実験条件

雑音	white, airp_j, bus_j,rest_j, offi_j, shop_j_1
雑音の長さ	3min
窓関数	ハニング窓
窓幅	512points (32ms)
シフト幅	256points (16ms)
周波数ビン	16〜64

A.3　結果

雑音のスペクトルコンポーネントの実部および虚部のヒストグラムをそれぞれ図 23〜図 30 に示した．

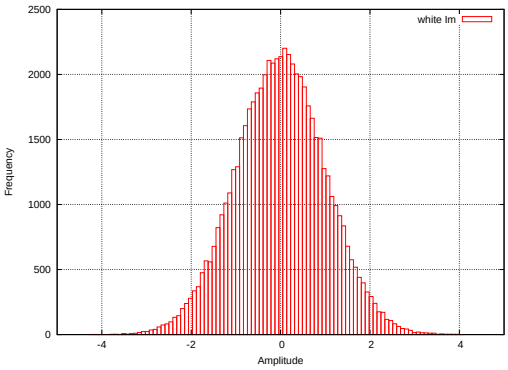

図 23: 白色雑音 (虚部)

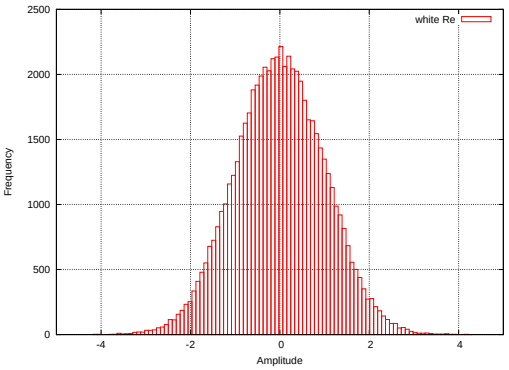

図 26: 白色雑音 (実部)

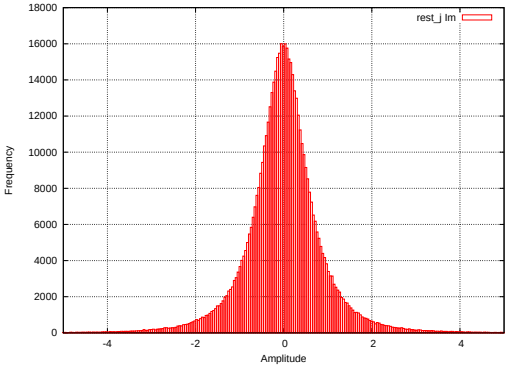

図 24: レストラン雑音 (虚部)

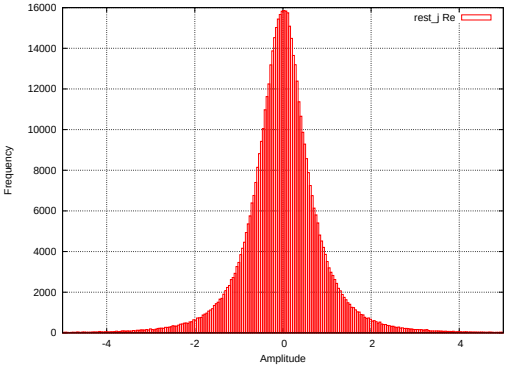

図 27: レストラン雑音 (実部)

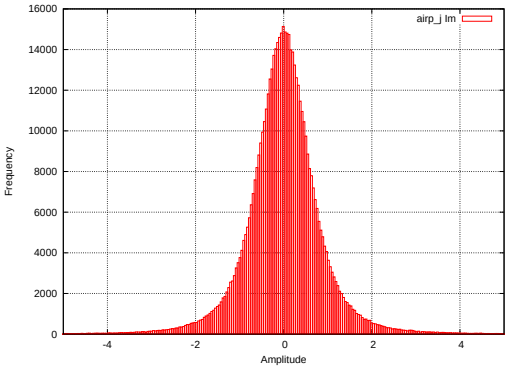

図 25: 空港雑音 (虚部)

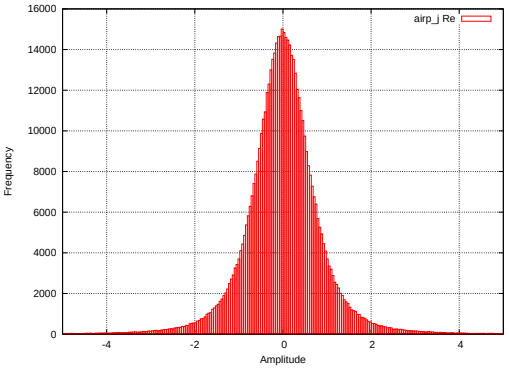

図 28: 空港雑音 (実部)

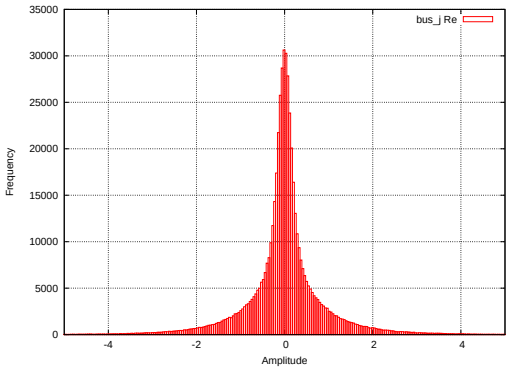

図 29: バス雑音 (実部)

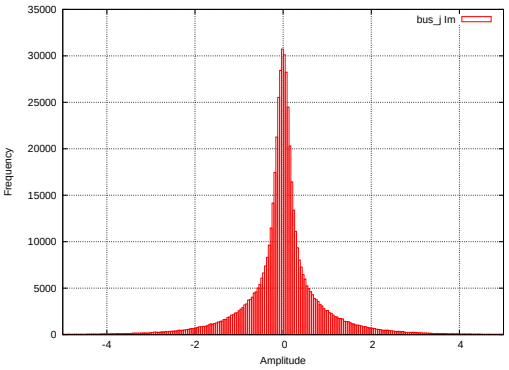

図 32: バス雑音 (虚部)

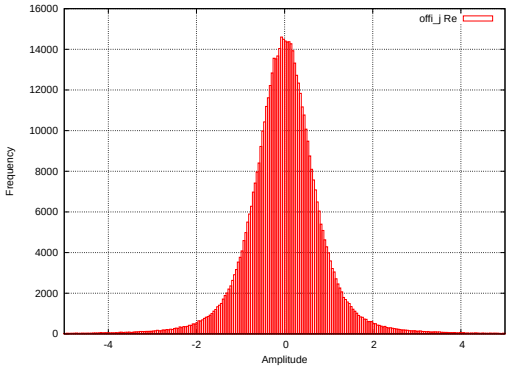

図 30: オフィス雑音 (実部)

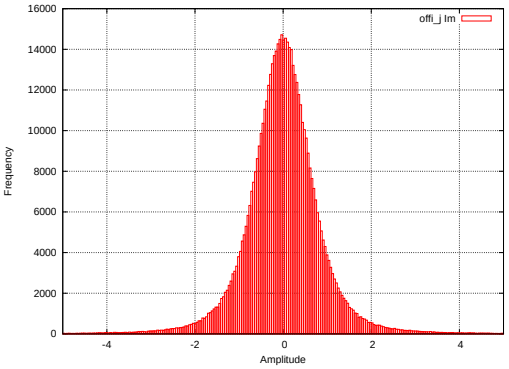

図 33: オフィス雑音 (実部)

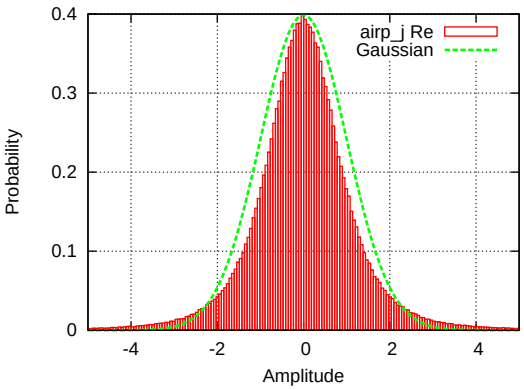

図 31: 空港雑音とガウス分布

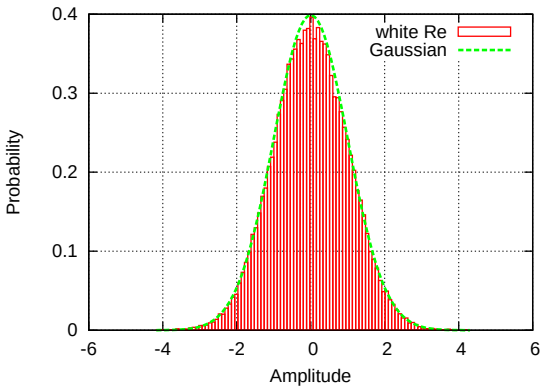

図 34: 白色雑音とガウス分布

A.4　考察

実部と虚部について　実部と虚部では分布の形状に違いは見受けられなかった．これは音声信号の場合と同様である．

分布の形状について　雑音の FFT コンポーネントはガウス分布に非常に近い形状をしていることが分かった．白色雑音の相対ヒストグラムはガウス分布にほぼぴったり一致した．しかし白色雑音以外の相対ヒストグラムでは，ガウス分布よりも尖った形状であった．空港雑音の相対ヒストグラムにガウス分布をフィッティングさせようとすると，小さい振幅の部分に合せようとすると大きい振幅の部分が合わなくなってしまった．逆に大きい振幅の部分に合せようとすると小さい振幅の部分が全く合わなくなってしまい，ガウス分布では表現しきれていないことが分った．スーパーガウシアンのほうがよりよく表現できると考えられる．

A.5　まとめ

本来のアンサンブルの特徴は知りえることができないので断言はできないが，ガウス分布では振幅が大きい部分について表現しきれていないことがわかった．ガウス分布の代りに $1/\cosh^2(x)$ という式で表現されるスーパーガウシアンや，一般化ガウス分布を用いることを考えたが，分布を完全には表現できないのではないかという懸念と，ゲイン関数の導出において他項式の 1/2 乗がでてきてしまうので，導出が非常に難しくなってしまうという問題がある．

そこで平均が音声で分散の値が異る 2 つのガウスを用いて観測信号の分布を表現する．これは別の分布を用いるよりも計算が簡単であることが利点になる．

B　位相が認識率に与える影響

本節では，クリーン音声の振幅と雑音重畳音声の位相を用いて合成した信号の音声認識率を調査し，音声信号の振幅スペクトルのみを推定する STSA 法の妥当性を調査した．

B.1　実験 1

式 (44) のようにクリーン音声の振幅と雑音重畳音声の位相を用いて合成した信号の単語正解率を調査した．なお時間領域の信号は，$\tilde{X}_k$ に逆フーリエ変換を施すことで得る．

$$\tilde{X}_k = A_k \exp(j\vartheta_k) \tag{44}$$

B.1.1　条件

実験条件を表 6 に示した．

表 4: 実験条件

SNR	0dB
窓関数	ハニング窓
窓幅	512points (32ms)
シフト幅	256points (16ms)
認識器	Viterbi デコーダ 3 状態, モノフォン
特徴量	MFCC
高域強調係数	0.97

B.1.2　結果

雑音ごとの単語正解率を図 35 に示し，単語正解率とそれぞれの雑音と CLEAN との単語正解率の差を表 5 に示した．

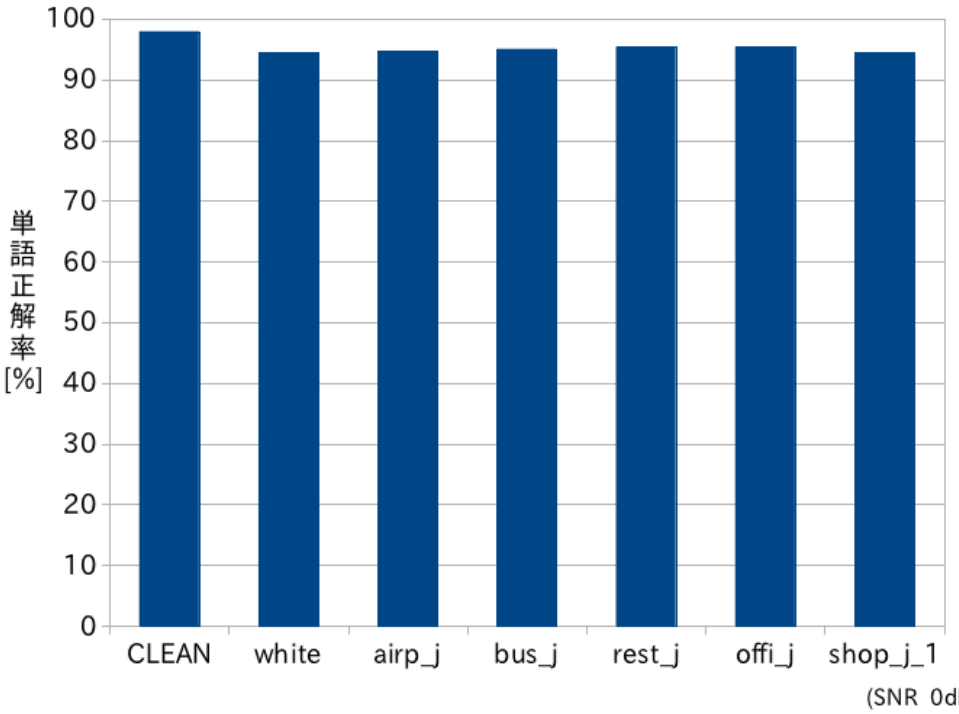

図 35: クリーン音声の振幅と雑音重畳音声の位相を用いて合成した信号の音声認識率

表 5: 単語正解率とそれぞれの雑音と CLEAN との単語正解率の差

	CLEAN	white	airp_j	bus_j	rest_j	offi_j	shop_j_1
単語正解率 [%]	97.90	94.48	94.62	95.02	95.37	95.37	94.38
CLEAN との差	0.00	3.42	3.28	2.88	2.53	2.53	3.52

B.2 実験 2

式 (45) のように観測信号から推定した音声の振幅スペクトルと，真の音声の位相を用いて合成した信号の単語正解率を調査し，位相成分の推定の必要性を調査した．なお時間領域の信号は，$\tilde{X}'_k$ に逆フーリエ変換を施すことで得る．

$$\hat{X}'_k = \hat{A}_k \exp(j\alpha_k) \tag{45}$$

B.2.1 条件

以下の 2 種類の信号を実験に用いる．その他の実験条件を表 6 に示した．

・信号 $\hat{X}$ $\cdots$ $\hat{X}_k = \hat{A}_k \exp(j\vartheta_k)$ で合成した信号

・信号 $\hat{X}'$ $\cdots$ $\hat{X}'_k = \hat{A}_k \exp(j\alpha_k)$ で合成した信号

表 6: 実験条件

音声	ATR set A
雑音	white, airp_j, bus_j, rest_j, offi_j, shop_j_1
SNR	0dB, 5dB, 10dB, 15dB
推定器	Parametric MMSE STSA [2]
音声の分布形状パラメータ	0.3
窓関数	ハニング窓
窓幅	512points (32ms)
シフト幅	256points (16ms)
認識器	Viterbi デコーダ 3 状態, モノフォン
特徴量	MFCC
高域強調係数	0.97

B.2.2 結果

信号 $\hat{X}$ と信号 $\hat{X}'$ の単語正解率を図 36 に示した．

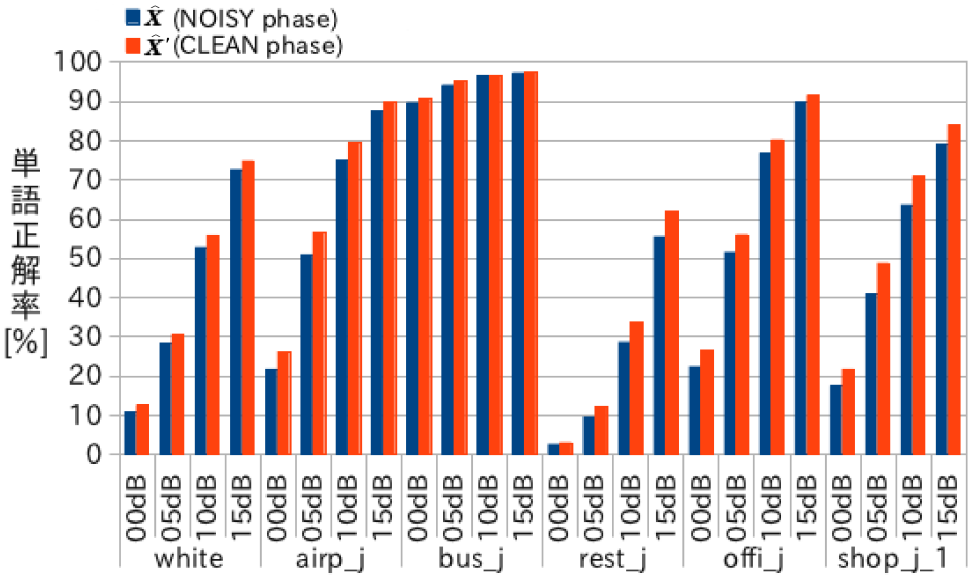

図 36: 単語正解率

B.3 まとめ

振幅成分が完璧に推定された場合には，位相成分の推定も正確に行うべきであると考えられる．また図 36 より，位相成分を正確に推定できれば単語正解率の向上は可能であることがわかった．しかし，表 5 から分るように雑音重畳音声の位相を用いたことによる単語正解率の悪化は最大で 3.52 ポイントであった．したがって，音声の振幅スペクトルのみを推定する STSA 法は妥当であると結論づける．

また本研究では，周波数領域で音声の振幅スペクトルの推定を行い，雑音重畳音声の位相を用いて一度時間領域に変換してから認識器に入力している．しかし多くの音声認識器には特徴量に MFCC を用いているので，振幅成分の推定を行った後に時間領域に変換せずに直接特徴量抽出を行えば位相成分の推定は考えないで良いこととなる．

やはり雑音低減の現状としては，位相成分の推定を行うよりも振幅スペクトルの推定精度を上げることが重要であると考えられる．

C　振幅と位相の独立性

本節では，振幅と位相の独立性を調査した結果を報告し，実際の信号で振幅と位相の独立性を仮定してよいことを検証した．

C.1　実験

位相に対する振幅の散布図を作成することで，振幅の位相による変化があるか否かを調査した．なお 0 番目のビンつまり直流成分は離散フーリエ変換の式からわかるように虚部が 0 であるので，見やすさのために取り除いて散布図を作成した．その他の周波数ビンは全て一緒にプロットした．

C.2　実験条件

実験条件を表 7 に示した．

表 7: 実験条件

使用データ	ATR 研究用データベースセット B (fknsda01, mhosda01)
窓関数	ハニング窓
窓幅	512points (32ms)
フレームシフト幅	256points (16ms)

C.3　結果と考察

fknsda01 の位相に対する振幅の散布図を図 37 に，mhosda01 の位相に対する振幅の散布図を図 38 に示した．

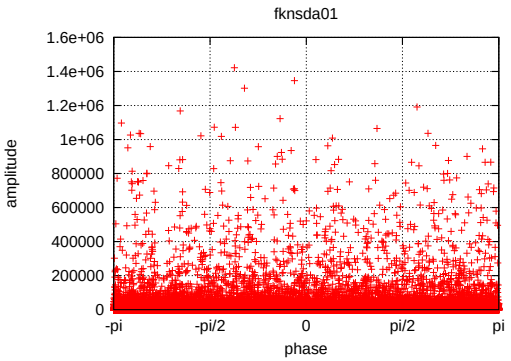

図 37: fknsda01 の位相に対する振幅

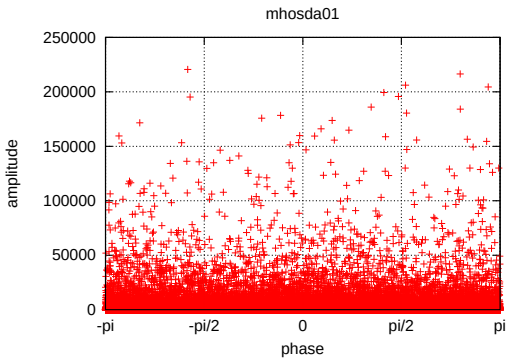

図 38: mhosda01 の位相に対する振幅

C.4　まとめ

散布図より位相は振幅に依存しないと考え，独立性を仮定して推定器の導出を行って問題ないと結論づける．

D　音声の DFT コンポーネントにガウス分布を仮定することと振幅にレイリー分布を仮定することの等価性の証明

音声のスペクトルコンポーネントにガウス分布を仮定することが振幅にレイリー分布を仮定することと等価であることを以下に示す.

まず本節では，図 39，式 (46)〜(46) のように $\mathrm{Re}\,[X]$ を x，$\mathrm{Im}\,[X]$ を y とする.

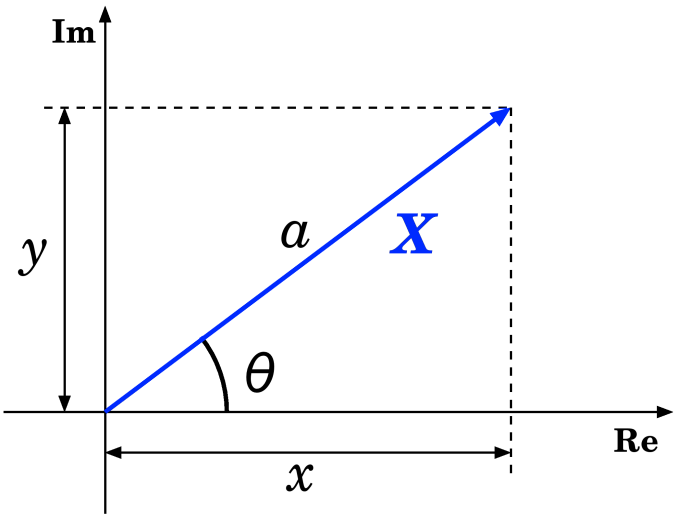

図 39: 本節における変数の関係

$$x = a\cos\theta \qquad (46) \qquad\qquad a = \sqrt{x^2 + y^2} \qquad (48)$$

$$y = a\sin\theta \qquad (47) \qquad\qquad \theta = \tan^{-1}\frac{y}{x} \qquad (49)$$

確率変数 x, y は互いに独立で，それぞれ平均 0 で分散 σ^2 のガウス分布に従うものとする.

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right) \qquad (50)$$

$$p(y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{y^2}{2\sigma^2}\right) \qquad (51)$$

すると x, y の結合確率密度関数 $p(x, y)$ は,

$$p(x, y) = p(x)\ p(y) \qquad (52)$$

と表すことができる．また，a, θ の結合確率密度関数 $p(a, \theta)$ はヤコビアン J により式 (53) のように表すことができる.

$$\begin{aligned} p(a, \theta) &= J\ p(x, y) \\ &= J\ p(x)\ p(y) \end{aligned} \qquad (53)$$

ただしヤコビアン J は以下とおりである．

$$\begin{aligned} J &= \begin{vmatrix} \frac{\partial x}{\partial a} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial a} & \frac{\partial y}{\partial \theta} \end{vmatrix} \\ &= \frac{\partial x}{\partial a}\frac{\partial y}{\partial \theta} - \frac{\partial x}{\partial \theta}\frac{\partial y}{\partial a} \\ &= (\cos\theta \ a\cos\theta) + (\sin\theta \ a\sin\theta) \\ &= a \ (\cos^2\theta + \sin^2\theta) \\ &= a \end{aligned} \tag{54}$$

したがって式 (53) は以下のように変形できる．

$$\begin{aligned} p(a,\theta) &= J \ p(x) \ p(y) \\ &= a \ \left\{ \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \right\} \ \left\{ \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right) \right\} \\ &= \frac{a}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \\ &= \frac{a}{2\pi\sigma^2} \exp\left(-\frac{a^2}{2\sigma^2}\right) \end{aligned} \tag{55}$$

ここで上式を θ に関して積分する．

$$\begin{aligned} p(a) &= \int_0^{2\pi} p(a,\theta) d\theta \\ &= \int_0^{2\pi} \frac{a}{2\pi\sigma^2} \exp\left(-\frac{a^2}{2\sigma^2}\right) d\theta \\ &= 2\pi \times \frac{a}{2\pi\sigma^2} \exp\left(-\frac{a^2}{2\sigma^2}\right) \\ &= \frac{a}{\sigma^2} \exp\left(-\frac{a^2}{2\sigma^2}\right) \end{aligned} \tag{56}$$

以上より，スペクトルコンポーネントの実部と虚部のそれぞれに平均 0, 分散 σ^2 のガウス分布を仮定することは，振幅スペクトルにレイリー分布を仮定することと等しいことを示した．

E Parametric MMSE STSA GM の導出

本節では，観測信号の確率密度関数には以前と同様に 2 混合の混合ガウス分布を用い，音声の確率密度関数に一般化ガンマ分布を用いた Parametric MMSE STSA GM 推定器の導出を行う．

$$\hat{A}_k = \frac{\int_0^\infty \int_0^{2\pi} a_k \; p(Y \mid a_k, \alpha_k) \; p(a_k, \alpha_k) \; d\alpha_k \; da_k}{\int_0^\infty \int_0^{2\pi} \quad p(Y \mid a_k, \alpha_k) \; p(a_k, \alpha_k) \; d\alpha_k \; da_k} \tag{57}$$

まず式 (57) の分子だけに着目して計算する．なお見易さのため，平均 X, 分散 λ のガウス分布を $N(X, \sqrt{\lambda})$ と置き，周波数ビン番号 k は中略した．

$$\begin{aligned}
(\text{分子}) &= \int_0^\infty \int_0^{2\pi} a \; p(Y \mid a, \alpha) \; p(a, \alpha) \; d\alpha \; da \\
&= \int_0^\infty \int_0^{2\pi} a \cdot \frac{1}{2\pi} \cdot \frac{2\Phi^\eta}{\Gamma(\eta)} a^{2\eta-1} \exp\left(-\Phi a^2\right) \cdot \left\{ rN(X, \sqrt{z\lambda_d}) + (1-r)N(X, \sqrt{w\lambda_d}) \right\} \; d\alpha \; da \\
&= \int_0^\infty \int_0^{2\pi} a \cdot \frac{1}{2\pi} \cdot \frac{2}{\Gamma(\eta)} \left(\frac{\eta}{\lambda_x}\right)^\eta a^{2\eta-1} \exp\left(-\frac{\eta}{\lambda_x} a^2\right) \\
&\qquad\qquad \cdot \left\{ rN(X, \sqrt{z\lambda_d}) + (1-r)N(X, \sqrt{w\lambda_d}) \right\} \; d\alpha \; da \\
&= \frac{1}{\pi\lambda_d \Gamma(\eta)} \left(\frac{\eta}{\lambda_x}\right)^\eta \int_0^\infty \int_0^{2\pi} a^{2\eta} \cdot \exp\left(-\frac{\eta}{\lambda_x} a^2\right) \cdot \left\{ rN(X, \sqrt{z\lambda_d}) + (1-r)N(X, \sqrt{w\lambda_d}) \right\} \; d\alpha \; da \\
&= \frac{1}{\pi\lambda_d \Gamma(\eta)} \left(\frac{\eta}{\lambda_x}\right)^\eta \left\{ \frac{r}{z} \int_0^\infty a^{2\eta} \exp\left(-\frac{\eta}{\lambda_x} a^2\right) \int_0^{2\pi} \exp\left(-\frac{R^2 + a^2 - 2Ra\cos(\vartheta - \alpha)}{z\lambda_d}\right) \; d\alpha \; da \right. \\
&\qquad \left. + \frac{r-1}{w} \int_0^\infty a^{2\eta} \exp\left(-\frac{\eta}{\lambda_x} a^2\right) \int_0^{2\pi} \exp\left(-\frac{R^2 + a^2 - 2Ra\cos(\vartheta - \alpha)}{w\lambda_d}\right) \; d\alpha \; da \right\} \\
&= \frac{1}{\pi\lambda_d \Gamma(\eta)} \left(\frac{\eta}{\lambda_x}\right)^\eta \left\{ \frac{r}{z} \int_0^\infty a^{2\eta} \exp\left(-\frac{\eta}{\lambda_x} a^2\right) \exp\left(-\frac{R^2 + a^2}{z\lambda_d}\right) \int_0^{2\pi} \exp\left(\frac{2Ra}{z\lambda_d}\cos(\vartheta - \alpha)\right) \; d\alpha \; da \right. \\
&\qquad \left. + \frac{r-1}{w} \int_0^\infty a^{2\eta} \exp\left(-\frac{\eta}{\lambda_x} a^2\right) \exp\left(-\frac{R^2 + a^2}{w\lambda_d}\right) \int_0^{2\pi} \exp\left(\frac{2Ra}{w\lambda_d}\cos(\vartheta - \alpha)\right) \; d\alpha \; da \right\}
\end{aligned} \tag{58}$$

ここで数式表 [17, eq. 8.431.5]

$$I_n(z) = \frac{1}{2\pi} \int_0^{2\pi} \cos(\beta n) \exp(z\cos\beta) d\beta$$

に $n = 0$ を代入して変形すると，

$$\int_0^{2\pi} \exp(z\cos\beta) d\beta = 2\pi I_0(z)$$

が成り立つ．また数式表 [17, eq. 8.406.3]

$$I_n(z) = i^{-n} J_n(iz) \tag{59}$$

に $n = 0$ を代入すると，

$$I_0(z) = J_0(iz) \tag{60}$$

が成り立つ．ただし本節における i は，虚数単位を意味する．また，$\frac{R}{z\lambda_d}$ を以下のように変形する．

$$\begin{aligned}
\frac{R}{z\lambda_d} &= \left(\frac{R^2}{z^2\lambda_d^2}\right)^{1/2} \\
&= \left(\frac{R^2}{z^2\lambda_d^2}\ \frac{\frac{1}{\eta}\lambda_x}{\frac{1}{\eta}\lambda_x + z\lambda_d}\ \frac{\frac{1}{\eta}\lambda_x + z\lambda_d}{\frac{1}{\eta}\lambda_x}\right)^{1/2} \\
&= \left(\frac{R^2}{z\lambda_d}\ \frac{\lambda_x/\lambda_d}{\lambda_x/\lambda_d + z\eta}\ \frac{\frac{1}{\eta}\lambda_x + z\lambda_d}{\frac{1}{\eta}\lambda_x\ z\lambda_d}\right)^{1/2} \\
&= \left(\frac{\xi}{\xi + z\eta}\ \frac{\gamma}{z}\ \frac{\frac{1}{\eta}\lambda_x + z\lambda_d}{\frac{1}{\eta}\lambda_x\ z\lambda_d}\right)^{1/2} \\
&= \left\{\frac{\xi}{\xi+1}\gamma_z\left(\frac{1}{t\lambda_x} + \frac{1}{z\lambda_d}\right)\right\}^{1/2} \\
&= \sqrt{\frac{v(\eta, z)}{\lambda(\eta, z)}}
\end{aligned} \tag{61}$$

ただし

$$v(\eta, \varsigma) \triangleq \frac{\xi}{\varsigma\eta + \xi}\frac{\gamma}{\varsigma} \tag{62}$$

$$\frac{1}{\lambda(\eta, \varsigma)} \triangleq \frac{\eta}{\lambda_x} + \frac{1}{\varsigma\lambda_d} \tag{63}$$

とする．

式 (59) と式 (60), 式 (61), を用いると，式 (58) 中の α に関する積分は以下のようになる．

$$\begin{aligned}
\int_0^{2\pi} \exp\left\{\frac{2Ra}{z\lambda_d}\cos(\vartheta - \alpha)\right\} d\alpha &= 2\pi I_0\left(\frac{2R}{z\lambda_d}a\right) \\
&= 2\pi J_0\left(i\frac{2R}{z\lambda_d}a\right) \\
&= 2\pi J_0\left(i2\frac{R}{z\lambda_d}a\right) \\
&= 2\pi J_0\left(i2\sqrt{\frac{v(\eta, z)}{\lambda(\eta, z)}}\ a\right)
\end{aligned} \tag{64}$$

と変形できる．w についても同様である．

式 (64) を用いて式 (57) の分子を計算すると以下のように変形できる．

$$
\begin{aligned}
(\text{分子}) &= \frac{1}{\pi\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^{\eta}\left\{\frac{r}{z}\int_0^{\infty} a^{2\eta}\exp\left(-\frac{\eta}{\lambda_x}a^2\right)\exp\left(-\frac{R^2+a^2}{z\lambda_d}\right)\int_0^{2\pi}\exp\left(\frac{2Ra}{z\lambda_d}\cos(\vartheta-\alpha)\right)\,d\alpha\,da\right. \\
&\qquad\left.+\frac{r-1}{w}\int_0^{\infty} a^{2\eta}\exp\left(-\frac{\eta}{\lambda_x}a^2\right)\exp\left(-\frac{R^2+a^2}{w\lambda_d}\right)\int_0^{2\pi}\exp\left(\frac{2Ra}{w\lambda_d}\cos(\vartheta-\alpha)\right)\,d\alpha\,da\right\} \\
&= \frac{1}{\pi\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^{\eta}\left\{\frac{r}{z}\int_0^{\infty} a^{2\eta}\exp\left(-\frac{\eta}{\lambda_x}a^2\right)\exp\left(-\frac{R^2+a^2}{z\lambda_d}\right)2\pi J_0\left(2i\sqrt{\frac{v(\eta,z)}{\lambda(\eta,z)}}\,a\right)\,da\right. \\
&\qquad\left.+\frac{r-1}{w}\int_0^{\infty} a^{2\eta}\exp\left(-\frac{\eta}{\lambda_x}a^2\right)\exp\left(-\frac{R^2+a^2}{w\lambda_d}\right)2\pi J_0\left(2i\sqrt{\frac{v(\eta,w)}{\lambda(\eta,w)}}\,a\right)\,da\right\} \\
&= \frac{2\pi}{\pi\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^{\eta}\left\{\frac{r}{z}\exp\left(-\frac{R^2}{z\lambda_d}\right)\int_0^{\infty} a^{2\eta}\exp\left(-\frac{\eta}{\lambda_x}a^2-\frac{a^2}{z\lambda_d}\right)J_0\left(2i\sqrt{\frac{v(\eta,z)}{\lambda(\eta,z)}}\,a\right)\,da\right. \\
&\qquad\left.+\frac{r-1}{w}\exp\left(-\frac{R^2}{w\lambda_d}\right)\int_0^{\infty} a^{2\eta}\exp\left(-\frac{\eta}{\lambda_x}a^2-\frac{a^2}{w\lambda_d}\right)J_0\left(2i\sqrt{\frac{v(\eta,w)}{\lambda(\eta,w)}}\,a\right)\,da\right\} \\
&= \frac{2}{\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^{\eta}\frac{1}{\pi\lambda_d}\left\{\frac{r}{z}\exp\left(-\frac{R^2}{z\lambda_d}\right)\int_0^{\infty} a^{2\eta}\exp\left(-\frac{a^2}{\lambda(\eta,z)}\right)J_0\left(2i\sqrt{\frac{v(\eta,z)}{\lambda(\eta,z)}}\,a\right)\,da\right. \\
&\qquad\left.+\frac{1-r}{w}\exp\left(-\frac{R^2}{w\lambda_d}\right)\int_0^{\infty} a^{2\eta}\exp\left(-\frac{a^2}{\lambda(\eta,w)}\right)J_0\left(2i\sqrt{\frac{v(\eta,w)}{\lambda(\eta,w)}}\,a\right)da\right\} \\
&= \frac{2}{\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^{\eta}\left\{\frac{r}{z}\exp\left(-\frac{R^2}{z\lambda_d}\right)\int_0^{\infty} a^{2\eta}\exp\left(-\frac{a^2}{\lambda(\eta,z)}\right)J_0\left(2i\sqrt{\frac{v(\eta,z)}{\lambda(\eta,z)}}\,a\right)\,da\right. \\
&\qquad\left.+\frac{1-r}{w}\exp\left(-\frac{R^2}{w\lambda_d}\right)\int_0^{\infty} a^{2\eta}\exp\left(-\frac{a^2}{\lambda(\eta,w)}\right)J_0\left(2i\sqrt{\frac{v(\eta,w)}{\lambda(\eta,w)}}\,a\right)da\right\}
\end{aligned}
\tag{65}
$$

さらに数式表 [17, 6.631] の

$$
\int_0^{\infty} x^{\mu}\exp\left\{-\alpha x^2\right\}J_{\nu}(\beta x)\,dx = \frac{\beta^{\nu}\Gamma(\frac{1}{2}\nu+\frac{1}{2}\mu+\frac{1}{2})}{2^{\nu+1}\alpha^{\frac{1}{2}(\mu+\nu+1)}\Gamma(\nu+1)}\,{}_1F_1\left(\frac{\nu+\mu+1}{2};\nu+1;-\frac{\beta^2}{4\alpha}\right) \tag{66}
$$

に $\nu=0, x=a, \mu=2\eta$ を代入すると，

$$
\begin{aligned}
\int_0^{\infty} a^{2\eta}\exp\left\{-\alpha a^2\right\}J_0(\beta a)\,da &= \frac{\beta^0\Gamma(\eta+\frac{1}{2})}{2\alpha^{(\eta+\frac{1}{2})}\Gamma(1)}\,{}_1F_1\left(\eta+0.5;1;-\frac{\beta^2}{4\alpha}\right) \\
&= \frac{\Gamma(\eta+\frac{1}{2})}{2\alpha^{(\eta+\frac{1}{2})}}\,{}_1F_1\left(\eta+0.5;1;-\frac{\beta^2}{4\alpha}\right)
\end{aligned}
\tag{67}
$$

が成り立つ．さらに上式に $\alpha = \frac{1}{\lambda(\eta,z)}, \beta = 2i\sqrt{\frac{v(\eta,z)}{\lambda(\eta,z)}}$ を代入する．

$$\begin{aligned}\int_0^\infty a^{2\eta}\exp\left\{-\alpha a^2\right\}J_0(\beta a)\ da &= \frac{\Gamma(\eta+\frac{1}{2})}{2\alpha^{(\eta+\frac{1}{2})}}\ {}_1F_1\left(\eta+0.5;1;-\frac{\beta^2}{4\alpha}\right)\\ &= \frac{1}{2}\Gamma\left(\eta+\frac{1}{2}\right)\lambda(\eta,z)^{(\eta+\frac{1}{2})}\ {}_1F_1\left(\eta+0.5;1;v(\eta,z)\right)\end{aligned} \tag{68}$$

式 (68) を用いて，式 (65) を次のように変形する．

$$\begin{aligned}(\text{分子}) &= \frac{2}{\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^\eta\left\{\frac{r}{z}\exp\left(-\frac{R^2}{z\lambda_d}\right)\int_0^\infty a^{2\eta}\exp\left(-\frac{a^2}{\lambda(\eta,z)}\right)J_0\left(2i\sqrt{\frac{v(\eta,z)}{\lambda(\eta,z)}}\ a\right)\ da\right.\\ &\qquad\left.+\frac{1-r}{w}\exp\left(-\frac{R^2}{w\lambda_d}\right)\int_0^\infty a^{2\eta}\exp\left(-\frac{a^2}{\lambda(\eta,w)}\right)J_0\left(2i\sqrt{\frac{v(\eta,w)}{\lambda(\eta,w)}}\ a\right)da\right\}\\ &= \frac{2}{\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^\eta\left\{\frac{r}{z}\exp\left(-\frac{R^2}{z\lambda_d}\right)\frac{1}{2}\ \Gamma\left(\eta+\frac{1}{2}\right)\lambda(\eta,z)^{(\eta+\frac{1}{2})}\ {}_1F_1\left(\eta+0.5;1;v(\eta,z)\right)\right.\\ &\qquad\left.+\frac{1-r}{w}\exp\left(-\frac{R^2}{w\lambda_d}\right)\frac{1}{2}\ \Gamma\left(\eta+\frac{1}{2}\right)\lambda(\eta,w)^{(\eta+\frac{1}{2})}\ {}_1F_1\left(\eta+0.5;1;v(\eta,w)\right)\right\}\\ &= \frac{2}{2\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^\eta\left\{\frac{r}{z}\exp\left(-\frac{R^2}{z\lambda_d}\right)\Gamma\left(\eta+\frac{1}{2}\right)\lambda(\eta,z)^{(\eta+\frac{1}{2})}\ {}_1F_1\left(\eta+0.5;1;v(\eta,z)\right)\right.\\ &\qquad\left.+\frac{1-r}{w}\exp\left(-\frac{R^2}{w\lambda_d}\right)\Gamma\left(\eta+\frac{1}{2}\right)\lambda(\eta,w)^{(\eta+\frac{1}{2})}\ {}_1F_1\left(\eta+0.5;1;v(\eta,w)\right)\right\}\\ &= \frac{1}{\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^\eta\left\{\frac{r}{z}\exp\left(-\frac{R^2}{z\lambda_d}\right)\Gamma\left(\eta+\frac{1}{2}\right)\lambda(\eta,z)^{(\eta+\frac{1}{2})}\ {}_1F_1\left(\eta+0.5;1;v(\eta,z)\right)\right.\\ &\qquad\left.+\frac{1-r}{w}\exp\left(-\frac{R^2}{w\lambda_d}\right)\Gamma\left(\eta+\frac{1}{2}\right)\lambda(\eta,w)^{(\eta+\frac{1}{2})}\ {}_1F_1\left(\eta+0.5;1;v(\eta,w)\right)\right\}\end{aligned} \tag{69}$$

式 (57) の分母についても分子の場合と同様に考えると以下の式が得られる．

$$\begin{aligned}(\text{分母}) = \frac{1}{\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^\eta&\left\{\frac{r}{z}\exp\left(-\frac{R^2}{z\lambda_d}\right)\Gamma(\eta)\,\lambda(\eta,z)^\eta\ {}_1F_1\left(\eta;1;v(\eta,z)\right)\right.\\ &\left.+\frac{1-r}{w}\exp\left(-\frac{R^2}{w\lambda_d}\right)\Gamma(\eta)\,\lambda(\eta,w)^\eta\ {}_1F_1\left(\eta;1;v(\eta,w)\right)\right\}\end{aligned} \tag{70}$$

以上の分子分母の変形と数式表により，式 (57) は以下のように変形できる．

$$\begin{aligned}\hat{A} = & \frac{\frac{1}{\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^{\eta}\Gamma\left(\frac{1}{2}+\eta\right)}{\frac{1}{\lambda_d\Gamma(\eta)}\left(\frac{\eta}{\lambda_x}\right)^{\eta}\Gamma(\eta)} \\ & \cdot \frac{\left\{\frac{r}{z}\ \lambda(\eta,z)^{\left(\frac{1}{2}+\eta\right)}\exp\left(-\frac{R^2}{z\lambda_d}\right)\ {}_1F_1\left(\frac{1}{2}+\eta;1;v(\eta,z)\right)\right.}{\left\{\frac{r}{z}\ \lambda(\eta,z)^{\eta}\ \exp\left(-\frac{R^2}{z\lambda_d}\right)\ {}_1F_1\left(\ \eta\ ;1;v(\eta,z)\right)\right.} \\ & \cdots \frac{\left.+\frac{1-r}{w}\ \lambda(\eta,w)^{\left(\frac{1}{2}+\eta\right)}\exp\left(-\frac{R^2}{w\lambda_d}\right)\ {}_1F_1\left(\frac{1}{2}+\eta;1;v(\eta,w)\right)\right\}}{\left.+\frac{1-r}{w}\ \lambda(\eta,w)^{\eta}\ \exp\left(-\frac{R^2}{w\lambda_d}\right)\ {}_1F_1\left(\ \eta\ ;1;v(\eta,w)\right)\right\}} \end{aligned} \tag{71}$$

$$\begin{aligned} = & \frac{\Gamma\left(\frac{1}{2}+\eta\right)}{\Gamma(\eta)} \\ & \cdot \frac{\left\{\frac{r}{z}\ \lambda(\eta,z)^{\left(\frac{1}{2}+\eta\right)}\exp\left(-\frac{R^2}{z\lambda_d}\right)\ {}_1F_1\left(\frac{1}{2}+\eta;1;v(\eta,z)\right)\right.}{\left\{\frac{r}{z}\ \lambda(\eta,z)^{\eta}\ \exp\left(-\frac{R^2}{z\lambda_d}\right)\ {}_1F_1\left(\ \eta\ ;1;v(\eta,z)\right)\right.} \\ & \cdots \frac{\left.+\frac{1-r}{w}\ \lambda(\eta,w)^{\left(\frac{1}{2}+\eta\right)}\exp\left(-\frac{R^2}{w\lambda_d}\right)\ {}_1F_1\left(\frac{1}{2}+\eta;1;v(\eta,w)\right)\right\}}{\left.+\frac{1-r}{w}\ \lambda(\eta,w)^{\eta}\ \exp\left(-\frac{R^2}{w\lambda_d}\right)\ {}_1F_1\left(\ \eta\ ;1;v(\eta,w)\right)\right\}} \end{aligned} \tag{72}$$

式 (72) の周波数ビン k に依存する変数に k を付与して変形すると式 (73) のようになる．

$$\hat{A}_k^{\mathrm{GM}} = \frac{r\cdot f_k(z,0.5)+(1-r)\cdot f_k(w,0.5)}{r\cdot f_k(z,0)+(1-r)\cdot f_k(w,0)} \tag{73}$$

$$f_k(\varsigma,\beta) = \frac{1}{\varsigma}\,\Gamma(\beta+\eta)\,\{\lambda_k(\eta,\varsigma)\}^{(\beta+\eta)}\cdot\exp\left(-\frac{R_k^2}{\varsigma\lambda_d(k)}\right)\cdot\ {}_1F_1\left(\beta+\eta;1;v_k(\eta,\varsigma)\right) \tag{74}$$

ただし，

$$v_k(\eta,\varsigma) \triangleq \frac{\xi_k}{\eta\varsigma+\xi_k}\frac{\gamma_k}{\varsigma} \tag{75}$$

$$\frac{1}{\lambda_k(\eta,\varsigma)} \triangleq \frac{\eta}{\lambda_x(k)}+\frac{1}{\varsigma\lambda_d(k)} \tag{76}$$

である．

以上より Parametric MMSE STSA GM 推定器の導出が示された。

F　分布形状パラメータ r に関する制約条件の導出

本節では，式 (29) の導出を行う．

まず平均 0 分散 λ_d のガウス分布を $N\left(0, \sqrt{\lambda_d}\right)$ と置き，雑音の分布にこのガウス分布 $N\left(0, \sqrt{\lambda_d}\right)$ を仮定する．このとき雑音の 2 次のモーメントは，

$$\begin{aligned} E\{\mid D \mid^2\} &= \int \mid D \mid^2 p(D) dD \\ &= \int \mid D \mid^2 N(0, \sqrt{\lambda_d}) dD \\ &= \lambda_d \end{aligned} \tag{77}$$

となる．

つぎに雑音に式 (78) の混合ガウス分布を仮定する．

$$p(D) = r \cdot N\left(0, \sqrt{z\lambda_d}\right) + (1-r) \cdot N\left(0, \sqrt{w\lambda_d}\right) \tag{78}$$

このとき雑音の 2 次のモーメントを式 (78) の混合ガウス分布を用いて計算すると，

$$\begin{aligned} E\{\mid D \mid^2\} &= \int \mid D \mid^2 \ p(D) \ dD \\ &= \int \mid D \mid^2 \left\{ r \cdot N\left(0, \sqrt{z\lambda_d}\right) + (1-r) \cdot N\left(0, \sqrt{w\lambda_d}\right) \right\} dD \\ &= r \cdot \int \mid D \mid^2 \ N\left(0, \sqrt{z\lambda_d}\right) \ dD + (1-r) \cdot \int \mid D \mid^2 \ N\left(0, \sqrt{w\lambda_d}\right) \ dD \\ &= r \cdot z \cdot \lambda_d + (1-r) \cdot w \cdot \lambda_d \end{aligned} \tag{79}$$

となる．

式 (77) と式 (79) より次の関係式が得られる．

$$\lambda_d = r \cdot z \cdot \lambda_d + (1-r) \cdot w \cdot \lambda_d \tag{80}$$

上式を r に関して整理すると以下のようになる．

$$r = \frac{1-w}{z-w} \tag{81}$$

以上より，式 (29) の導出が示された．